Molecular Genetics and Metabolism **68,** 103–125 (1999)
Article ID mgme.1999.2922, available online at http://www.idealibrary.com on IDEAL®

MINIREVIEW

The Structural Basis of Phenylketonuria

Heidi Erlandsen and Raymond C. Stevens[1]

The Scripps Research Institute, Department of Molecular Biology and Institute for Childhood and Neglected Diseases, La Jolla, California 92037

Received August 10, 1999

The human phenylalanine hydroxylase gene (*PAH*) (locus on human chromosome 12q24.1) contains the expressed nucleotide sequence which encodes the hepatic enzyme phenylalanine hydroxylase (PheOH). The PheOH enzyme hydroxylates the essential amino acid L-phenylalanine resulting in another amino acid, tyrosine. This is the major pathway for catabolizing dietary L-phenylalanine and accounts for approximately 75% of the disposal of this amino acid. The autosomal recessive disease phenylketonuria (PKU) is the result of a deficiency of PheOH enzymatic activity due to mutations in the *PAH* gene. Of the mutant alleles that cause hyperphenylalaninemia or PKU 99% map to the *PAH* gene. The remaining 1% maps to several genes that encode enzymes involved in the biosynthesis or regeneration of the cofactor ((*6R*)-L-erythro-5,6,7,8-tetrahydrobiopterin) regenerating the cofactor (tetrahydrobiopterin) necessary for the hydroxylation reaction. The recently solved crystal structures of human phenylalanine hydroxylase provide a structural scaffold for explaining the effects of some of the mutations in the *PAH* gene and suggest future biochemical studies that may increase our understanding of the PKU mutations.

Key Words: **phenylalanine hydroxylase; hyperphenylalaninemia; phenylketonuria; structure-function relationship.**

Phenylalanine hydroxylase (PheOH,[2] phenylalanine 4-monooxygenase, EC 1.14.16.1) is an iron-containing enzyme that catalyzes the conversion of the essential amino acid L-phenylalanine (L-Phe) into L-tyrosine, utilizing the cofactor 6(*R*)-L-erythro-tetrahydrobiopterin (BH_4) and dixoygen (for reviews, see (1–3)). This catabolic pathway accounts for approximately 75% or more of the disposal of dietary phenylalanine (4).

Mutations in the human gene encoding the PheOH enzyme (*PAH;* GenBank cDNA Reference Sequence U49897) result in the autosomal recessively inherited disease hyperphenylalaninemia (HPA). The resulting clinical phenotypes can range in severity from mild forms of non-PKU HPA, with benign outcome, to the severe form, phenylketonuria (PKU[3]; OMIM Entry 261600). Impaired function of PheOH results in the accumulation of high levels of blood plasma L-Phe and mental impairment if not treated with diet at an early age. Most of the mutant alleles map to the *PAH* gene, but approximately 1% map to several genes for enzymes involved in the biosynthesis or regeneration of cofactor (BH_4) (5). Due to an extensive and widely applied population screening worldwide for neonatal HPA, more than 380 mutations have been identified (including missense, nonsense, deletion, insertion, polymorphisms, and splice mutations) in the *PAH* gene (PAH Mutation Analysis Consortium database

[1] To whom correspondence should be addressed. Fax: (858) 784-9483. E-mail: stevens@scripps.edu.

[2] Abbreviations used: PKU; phenylketonuria, PheOH; phenylalanine hydroxylase, HPA; hyperphenylalaninemia; *PAH:* phenylalanine hydroxylase-encoding gene, BH_4; 6(*R*)-L-erythro-tetrahydrobiopterin (natural cofactor), 6-MPH_4; 6-methyltetrahydropterin, TnT; T7-coupled transcription-translation expression system.

[3] The PKU database is accessible at the PAH Mutation Analysis Consortium database and web site, located at http://www.mcgill.ca/pahdb (6–8).

1096-7192/99 $30.00

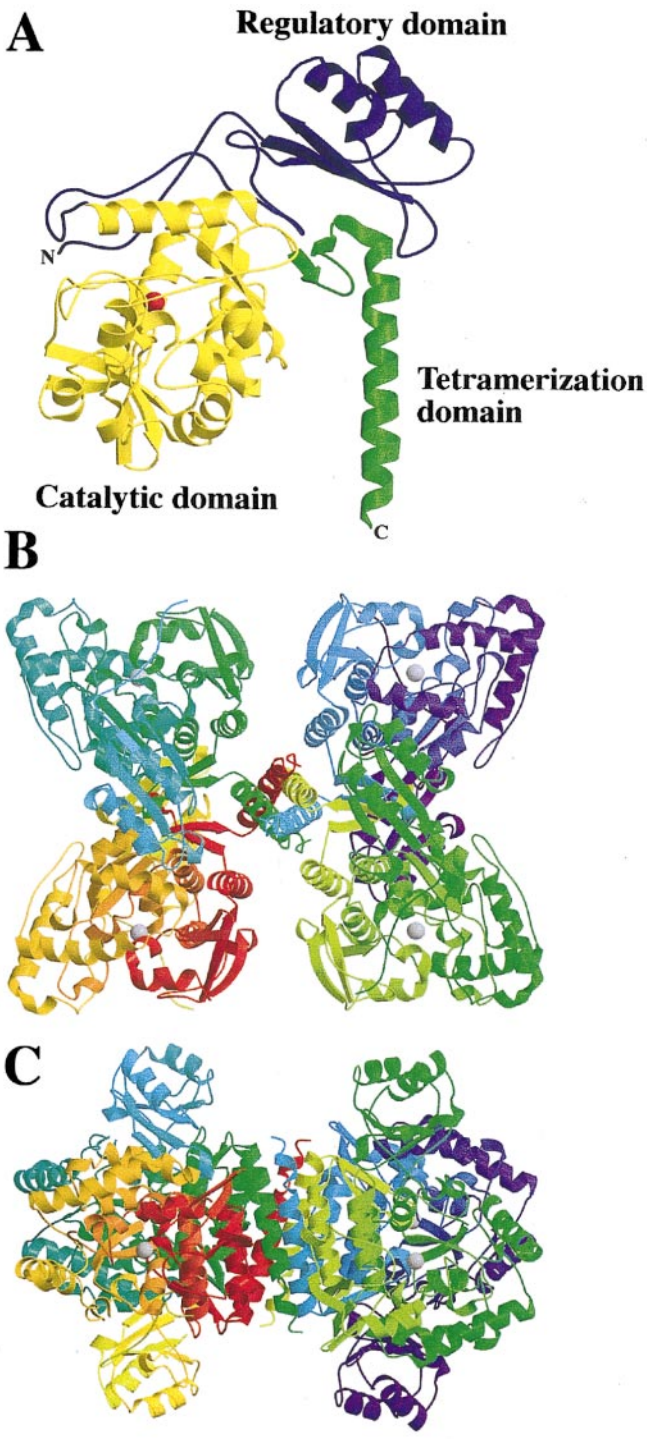


FIG. 1. (A) Structure of human phenylalanine hydroxylase full-length composite model. The model was generated using the structures of catalytic/tetramerization domains (9,10) and regulatory/catalytic domains (11) and superimposing the catalytic domains of the two models. The regulatory domain is shown in purple, the catalytic domain in gold, and the tetramerization domain in green. The iron is shown as a red sphere. (B, C) Two perpendicular views of the full-length PheOH model structure, colored from red (N-terminus in monomer A) to blue (C-terminus in monomer D). The iron is shown as a gray sphere in all four monomers.

and web site at http://www.mcgill.ca/pahdb, Refs. 6–8). Work on *in vitro* expression analysis of inherited human mutations is being carried out in order to link the genotype to the phenotype (4). This work can confirm that individual mutations impair PheOH enzyme activity and to a certain degree predict the severity of the corresponding metabolic phenotype. In this context, the recently solved structures of several truncated forms of PheOH (9–11) as well as the recent structure of the homologous enzyme tyrosine hydroxylase (TyrOH, tyrosine 3-monooxygenase, EC 1.14.16.2) with bound cofactor analogue (7,8-dihydrobiopterin) (12) will prove to be invaluable in the identification of amino acid residues involved in catalysis and regulation of enzymatic activity, as well as making it possible to give a complete analysis of the current (Spring 1998, PKU Newsletter) mutations in the *PAH* gene. Thus, in this review, a structural interpretation of the known mutations will be compared with the known data on *in vitro* expression analyses, PKU phenotypes, and both predicted and observed frequency of mutation in the *PAH* gene in order to have a more complete picture of the disease-causing PKU mutations.

Structure of Human Phenylalanine Hydroxylase

Recombinant human liver phenylalanine hydroxylase is reported to exist in solution as a pH-dependent equilibrium between homotetramers and homodimers (13). The molecular mass of one subunit is approximately 50 kDa (between 50 and 53 kDa). It has an absolute requirement for ferrous iron to be active and utilizes the cofactors BH_4 and O_2 to perform the hydroxylation reaction. PheOH, like the two other aromatic amino acid hydroxylases (TyrOH and tryptophan hydroxylase (TrpOH, tryptophan 5-monooxygenase, EC 1.14.16.4)), consists of three domains: the regulatory domain (residues 1–142), the catalytic domain (residues 143–410), and a short tetramerization domain (residues 411–452). The regulatory domain has a common α-β sandwich motif commonly observed for other regulatory domains, consisting of a four-stranded antiparallel β-sheet flanked on one side by two short α-helices and on the other side by the catalytic domain (11). The active-site iron resides in an open and solvent accessible region in the catalytic domain, the bottom of a novel basket-like arrangement of 14 α-helices and 8 β-strands. The short tetrameriza-

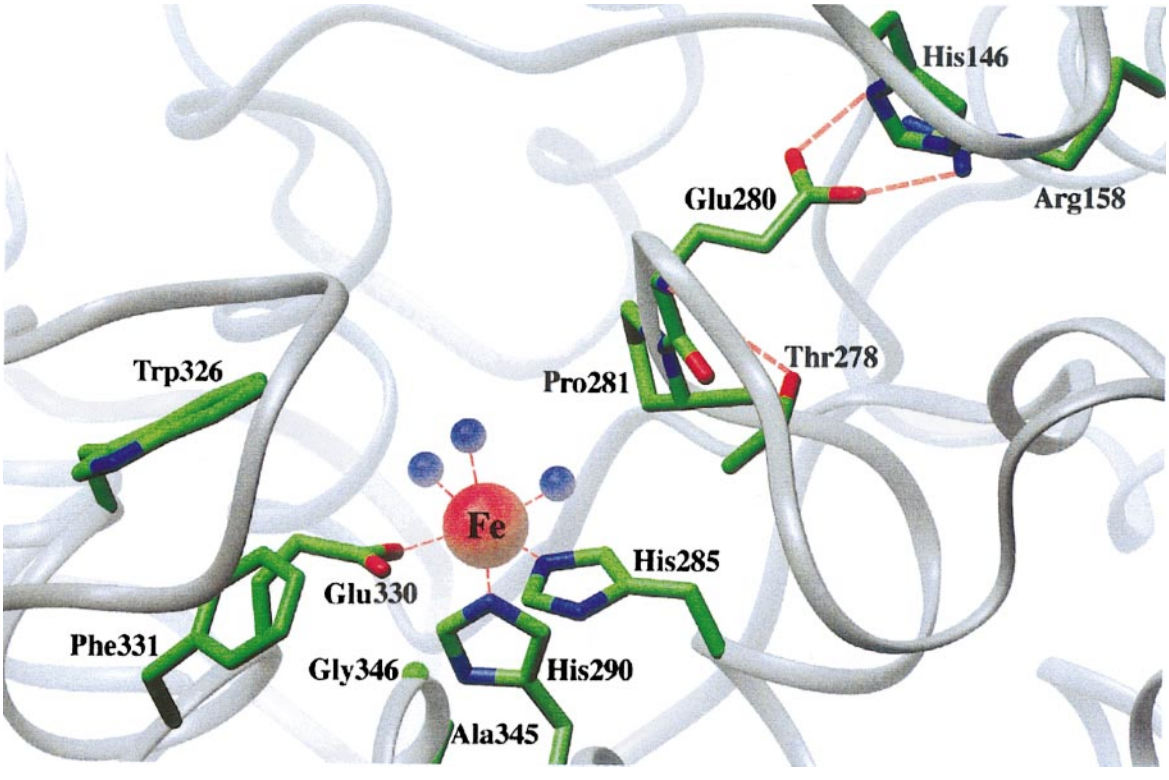


FIG. 2. Close-up of the active site surrounding the catalytic iron. The iron is shown as a red sphere and the three water molecules liganded to the iron are shown as blue spheres. Some residues that have reported PKU mutations (Thr278, Glu280, Pro281, Trp326, Phe331) and are located close to the iron in the active site, together with their interacting residues, are also shown.

tion domain consists of two antiparallel β-strands and a single 40 Å long C-terminal α-helix that is responsible for forming the domain-swapped coiled-coil core of the PheOH tetramer. Based on the crystal structures of several truncated forms of PheOH, including the regulatory/catalytic domains (11) and catalytic/tetramerization domains (9,10), a full-length model can be constructed by superimposing the respective catalytic domains (Fig. 1). The details of the secondary structure of the full-length model are shown in Fig. 2.

The active-site catalytic iron is ligated in an octahedral manner by two histidine residues (His285 and His290), one glutamic acid residue (Glu330), and three water molecules (Fig. 2). There also is a narrow tunnel connected to the active-site pocket, and this may be where the substrate is directed into the active site. A short loop made up of residues 378 to 381 forms one wall of the tunnel that runs out into the active-site pocket. This loop shows high mobility in the high-resolution structure of dimeric PheOH (9) and could be involved in restricting accessibility of substrates into the active site. The majority of the 34 amino acids lining the active site "bowl" are hydrophobic residues except for three charged glutamic acids, two putatively charged histidines, and an uncharged tyrosine.

PKU Mutations

Currently, most of the mutations that cause HPA or PKU have been characterized only at the DNA level. Recent genotype-phenotype correlation studies have revealed that the human PheOH genotype does not sufficiently explain the mutant phenotype in all cases, implying that other factors are involved in the emergent property of plasma homostasis (8). Some evidence toward understanding the molecular mechanisms leading to the phenotypic heterogeneity have been achieved through *in vitro* expression of recombinant forms of human PheOH (14). This method has proved to be important when it comes to characterization of the different mutations' individual catalytic, regulatory, and stability properties, mainly due to the very limited availability of endogeneous (wild-type, human liver) enzyme. These expression analyses, some in several *in vitro* expression systems, have revealed at least three groups of mutations that differ in their kinetic behavior and/or stability *in vivo:* (i) mutations affect-

TABLE 1
Single-Point Mutations Identified in the *PAH* Gene (not Including Silent and Missense Mutations)

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM*	Calc. FOM*	Reference
ATG/ATA	M1I	Met1 not visible in crystal structure solved			(28)
ATG/GTG	M1V		7.24	5.30	(29)
TCT/CCT	S16P	Ser16 not visible in crystal structure solved.	2.93		TC
TTC/TTG	F39L	π-stacks with Phe79 and Leu106.	3.16	7.07	(45)
TCA/TTA	S40L	Close to Tβ1 and Tβ2. In a charged region of RD. Substitution to Leu disturb electric potential.	7.96		(25)
CTC/TTC	L41F	Not room for a Phe. Close to Val51, Ala47 and Ile102.	2.94		(39)
AAA/ATA	K42I	Subst. to Ile disrupt Lys42 H-bond to Glu44. Stabilizes loop between Rβ1 and Rα1.	2.37	1.77	(46)
GGT/AGT	G46S	On surface of RD in loop at start of Rα1. H-bond partners available for a substitution to a Ser. Will result in distortion of secondary structure in RD.	7.12	35.35	(28)
GCA/GTA	A47V	Close to Leu41 and Lys42 at start of Rα1. Stabilizing start of helix. Difficult to accomodate	3.98	1.77	(47)
GCA/GAA	A47E	larger sidechains like Val and Glu.			TC
TTG/TCG	L48S	No predictable effect of substitution.	2.97	24.74	(48)
TTG/TCG	L52S	No predictable effect of substitution.	3.12		TC
CGC/TGC	R53C	No predictable effect of	14.9		TC
CGC/CAC	R53H	either substitution			TC
TTT/TTG	F55L	No predictable effect of substitution.	3.20		TC
GAG/GAT	E56D	No predictable effect of substitution.	5.48		(49)
GAT/TAT	D59Y	No predictable effect of substitution.	5.76		TC
AAC/AAG	N61K	H-bonds to His64 stabilizes loop between Rα1 and Rβ2.	2.39		TC
ACCCAC/	T63P/	Same as above. Substitution with Pro may distort	2.12/		(39)
CCCAAC	H64N	loop.	3.61		
ATT/GTT	I65V	In hydrophobic core of RD. Substitution to Val no predictable effect. Substitution to a more polar Thr			TC
ATT/ACT	I65T	or Asn may distort the hydrophobic packing in	4.72	38.89	(29)
ATT/AAT	I65N	the RD core.			(50)
TCT/CCT	S67P	In Rβ2, a substitution to a rigid Pro may distort secondary structure.	2.44		(36)
AGA/GGA	R68G	H-bonds to Ser67, and stabilizes secondary structure of Rβ2. Substitution to either Gly or Ser			TC
AGA/AGT	R68S	may disrupt this hydrogen bond.	4.72	1.77	(51)
TCT/CCT	S70P	H-bonds to Leu72 and Asn73. Stabilizes loop betwen Rβ2 and Rβ3.	2.11		TC
GAG/GCG	E76A	At start of Rβ3, Glu76 H-bonds to Asn73 and	5.26	1.77	TC
GAG/GGG	E76G	His208. A mutation to either Ala or Gly may distort secondary structure.			TC
GAA/AAA	E78K	Substitution to Lys would disrupt the stabilizing effect at Rβ3 of a saltbridge to Arg71.	5.45		TC
ACC/CCC	T81P	A substitution to a Pro disturb secondary structure of Rβ3.	2.72		(52)
GAT/TAT	D84Y	H-bond to Arg86 stabilizes start of α-helix Rα2. Substitution to Tyr destabilize PheOH structure.	4.54		(25)
AGC/AGA	S87R	Close to several hydrophobic residues. A substitution to an Arg could destabilize structure.	3.32	3.53	(47)
CCT/TCT	P89S	Pro89 forms a kink in middle of Rα2. Substitution	3.82		TC

TABLE 1—*Continued*

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM*	Calc. FOM&	Reference
		to a Ser distorts rigid secondary structure.			
ACA/ATA	T92I	On surface of protein. No predictable effect of substitution.	2.94	1.77	(39)
ATC/AGC	I94S	In hydrophobic pocket of RD. Substitution to a Ser may disturb hydrophobic packing.	4.43		TC
ATC/TTC	I95F	Close to two charged amino acids; Arg99 and Glu108. A substitution into a larger hydrophobic residue may not be easily accomodated.	4.43		TC
TTG/TCG	L98S	Same as for I94S.	3.36		(47)
GCC/GAC	A104D	A substitution to a larger and charged amino acid in loop between Rα2 and Rβ4 may destabilize the loop structure.	4.00	12.37	(30)
TCA/TTA	S110L	Hydrogen bonded to Lys85. Substitution to a Leu change/destabilize secondary structure.	7.14		TC
CCA/CAA	P122Q	In loop between Rβ5 and Cα1. May be important to start Cα1 helix in correct way. Substitution to a Gln distort start of helix Cα1.	3.20	1.77	(53)
ACC/ATC	T124I	Stabilizes start of Cα1 by H-bonding to Glu127. A substitution to a Ile would destabilize due to several polar residues being close to it.	2.10		(39)
GAC/TAC	D129Y	At end of Cα1, Asp 129 H-bonds to His170 and Gln127. Sustitutions to either Tyr or Gly will			TC
GAC/GGC	D129G	disrupt H-bonding and destabilize secondary structure.	4.25	1.77	(53)
GAT/GGT	D143G	On surface of entrance to active site. Possibly participates in controlling access to active site.	5.17		(17)
GAC/GTC	D145V	Same as above. On surface of active site "tunnel".	4.79		(25)
CAC/TAC	H146Y	A substitution to a Tyr would break a H-bonds to Pro147 at start of Cα2a and destabilize this secondary structure element	3.57		TC
CCT/TCT	P147S	A substitution to a Ser would remove a Pro which may be important for starting the Cα2a helix properly.	4.07	1.77	(53)
GGT/AGT	G148S	Close to two polar residues. May not be enough room to incorporate a Ser with proper secondary structure preserved.	7.12		(54)
GAT/CAT	D151H	Makes important stabilizing H-bonds to Pro152 and His271 in loop between Cα2a and Cα2b.			(36)
GAT/GGT	D151G	Substitution to either His or Gly destabilizing to secondary structure elements in area.	4.45	1.77	(53)
TAC/AAC	Y154N	H-bonds to Arg150 and Ile269 carbonyl-oxygen. π-stacks onto His271.	6.92		TC
AGA/AAC	R157N	Two salt bridges to Glu179 (in rat PheOH structure; Tyr in hPheOH) and Glu178. This holds the two helices Cα2b and Cα3 as well as the active site, together. These salt bridges would be disrupted in an Asn substituted mutant.	3.52		(30)
CGG/TGG	R158W	A substitution of Arg158 to a Trp or a Gln would disrupt an important salt-bridge to Glu280 and a			(56)
CGG/CAG	R158Q	H-bond to Tyr268, which is important for defining the shape of the active site.	20.60	21.21	(31,19)
CAG/CCG	Q160P	Substitution to a Pro in middle of Cα2 would disturb helix formation.	6.55		TC
TTT/TCT	F161S	Substitution into a Ser abolishes π-stacking onto Tyr179 in core of CD.	2.99		(55)
ATT/ACT	I164T	No predictable effect of substitution.	4.72	1.77	(24)
GCC/ACC	A165T	Substitution to any larger residue will	4.00		(53)
GCC/CCC	A165P	disturb the tight, internal packing in area.			TC

TABLE 1—*Continued*

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM[a]	Calc. FOM[a]	Reference
AAC/ATC	N167I	Last residue in Cα2b on the surface of CD. H-bonds to Asp 163 disrupted in an Ile mutation	2.79		(36)
TAC/CAC	Y168H	Subtitution into a His disrupts a H-bond to Gln172.	7.46		TC
CAT/CGT	H170R	Substitution into an Arg will disrupt a hydrogen bond to Arg241 at the start of Cβ1.	5.21	1.77	(28)
GGG/AGG	G171R	On surface.			(36)
GGG/GCG	G171A	No predictable effect of either subtitution	6.89	1.77	(24)
CCC/ACC	P173T	Pro173 in loop between Cα2b and Cα3. May be important for keeping helices together. Same effects as P175A mutation.	3.45		(57)
ATC/GTC	I174V	No predictable effect for either	4.30		TC
ATC/ACC	I174T	substitution.			(39)
CCT/GCT	P175A	H-bonds to Pro173 carbonyl-oxygen. Important together with Pro173 for positioning helices Cα2b and Cα3 correctly.	4.07		(25)
CGA/CCA	R176P	Any substitution would disrupt a			TC
CGA/CTA	R176L	salt-bridge to Asp229 in Cα5.	40.65	1.77	(47)
GTG/CTG	V177L	No predictable effect of substitution.	3.54		(28)
GAA/GGA	E178G	On surface of CD. Substitution to hydrophobic	4.86		(58)
GAA/GTA	E178V	residues very unfavourable?			TC
GAA/GGA	E182G	Same as mutation above.	4.81		(52)
TGG/CGG	W187R	In an aromatic patch of CD. Substitution into a	5.45		(59)
TGG/TGC	W187C	charged Arg or a polar Cys change hydrophobicity.			TC
GTG/GCG	V190A	No predictable effect of substitution.	3.54		TC
CTG/CCG	L194P	In middle of Cα3. Substitution into a Pro distorts helix.	3.83	1.77	(60)
CAT/CGT	H201R	H-bonds to Thr200. Important for loop structure.	5.21		(36)
CAT/TAT	H201Y	Substitution to Arg or Tyr disrupt H-bond.			(50)
TGC/TAC	C203Y	Substitution to Tyr disrupts a disulfide bridge formed to Cys334.	10.10		TC
TAT/TGT	Y204C	On surface of CD. No predictable effect of substitution.			(32)
TAC/GAC	Y206D	Close to several aromatic residues. Substitution to Asp disrupts a hydrogen bond to Leu288.	7.19		(58)
CCA/ACA	P211T	Pro211 makes a kink in helix Cα4. Substitution into Thr change helix structure.	2.90		(39)
CTT/CCT	L212P	In helix Cα4. Substitution to Pro disturbs helix structure.	3.60		(46)
CTT/CCT	L213P	Same as above.	3.21	1.77	(25)
TGT/GGT	C217G	No contacts. No predictable effect of substitution.	13.37		(61)
GGC/GTC	G218V	On surface of CD. No predictable effect of substitution.	6.17	5.30	(39)
GAA/GGA	E221G	On surface of CD. Substitution to a Gly disrupts a H-bond to Lys192, also on surface of CD.	5.40		(48)
GAT/GTT	D222V	Stabilization of loop between Cα4 and Cα5 by H-			(36)
GAT/GGT	D222G	bond to His220 disrupted if substituted to a Val.	4.39	1.77	TC
ATT/ACT	I224T	Close to Tyr179. No predictable effect of substitution to Thr.	4.61		(62)
ATT/ATG	I224M	No room for a Met?			(57)
CCC/ACC	P225T	Just before start of helix Cα5. May help	3.40		(63)
CCC/GCC	P225A	start helix. Any substitution to another			TC
CCC/CGC	P225R	residue changes helix.			TC
GTT/ATT	V230I	In hydrophobic patch of CD. No predictable effect of substitution.	6.48	7.07	(39)
TCT/CCT	S231P	A substitution to a Pro inside Cα5 disrupts a H-bond to Leu242 as well as distort helix.	2.11	3.54	(64)
TTC/TTA	F233L	No predictable effect of substitution.	2.88		(36)
ACT/CCT	T238P	Substitution to Pro disrupt H-bond to Phe294	2.73		(65)

TABLE 1—*Continued*

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM[a]	Calc. FOM[a]	Reference
		carbonyl-oxygen at end of Cα5.			
GGT/AGT	G239S	On surface - close to TD. No	7.12	1.77	(24)
GGT/GCT	G239A	predictable effect of substitutions.		1.77	TC
GGT/GTT	G239V				TC
TTC/GTC	F240V	Close to active site area as well as several Phe`s. π-stacks onto Phe294. Disrupted on subst. to Val.	2.46		TC
CGC/CAC	R241H	On surface of CD. No predictable effect			TC
CGC/CTC	R241L	of substitutions.	19.71	1.77	(66)
CGC/TGC	R241C			1.77	(67)
CTC/TTC	L242F	In center of CD, close to several hydrophobic residues. May not be room for a substituted Phe.	3.45		(65)
CGA/CAA	R243Q	Substitution to Gln disrupt salt-bridge to Asp129, and destabilize Cβ1 interaction with Cα1.	52.76	5.30	(32)
CCT/CTT	P244L	Last residue in Cβ1. Substitution to Leu destabilizes loop between Cβ1 and Cα6.	3.62	5.30	(68)
GTG/CTG	V245L				(25)
GTG/ATG	V245I	No predictable effect for any			TC
GTG/GAG	V245E	substitutions.			(39)
GTG/GCG	V245A		4.52	8.84	(47)
GCT/GAT	A246D	No predictable effect for either	3.89		(36)
GCT/GTT	A246V	substitutions.			TC
GGC/GTC	G247V	No predictable effect of substitution.	6.17		(55)
CTG/CGG	L248R	A substituted Arg will protrude into active site	3.88		TC
CTG/CCG	L248P	whereas a Pro will change shape of active site.			(61)
CTT/CAT	L249H	On surface of active site. Both substitutions can			TC
CTT/TTT	L249F	be accomodated, but will protrude further into active site than Leu.	3.90	1.77	(69)
CGG/CAG	R252Q	Forms a salt-bridge to Asp315 and a H-bond to the carbonyl- oxygen of Ala313 as well as Asp26.	21.25	1.77	(18)
CGG/GGG	R252G	Any substitution would destabilize the			(18)
CGG/TGG	R252W	structure of PheOH.		10.61	(23,20)
TTC/ATC	F254I	Located in active site (5.9Å away from Fe). Substitution to Ile distort important cofactor π-stacking interactions.	2.38		TC
TTG/GTG	L255V	Important for defining shape of active site.	3.32		(55,18)
TTG/TCG	L255S				(18)
GGC/TGC	G257C				TC
GGC/AGC	G257S	At start of Cα6. Close to Leu128. Any	6.17		TC
GGC/GAC	G257D	substitution disturbs secondary structure.			TC
GGC/GTC	G257V				TC
					(20,18)
GCC/ACC	A259T	In middle of Cα6. Close to Leu311. Any	3.73	1.77	
GCC/GTC	A259V	substitution would disturb secondary structure.			(18,20)
		In loop between Cα6 and Cβ2. Any substitution			(19,20)
CGA/CAA	R261Q	would disrupt H-bonds to Gln304 and Thr238	52.76	60.98	
CGA/CCA	R261P	which stabilizes the secondary structure in the active site.			TC
CGA/TGA	R261X	Premature stop in translation.		10.61	(70)
TTC/TTG	F263L	Is located 4.9Å away from Fe in active site. Possible π-stacking interactions with substrate may be impossible with substitution to a Leu.	3.24		(56)
CAC/CTC	H264L	No contacts made. No predictable effect of substitution.	3.62		TC
TGC/TAC	C265Y	Close to active site Fe. No space for substitution into a Tyr.	11.69		TC
TGC/GGC	C265G	No predictable effect of substitution into a Gly.			TC
TAC/CAC	Y268H	H-bond to Pro281 carbonyl-oxygen possibly broken if substituted to a His.	7.19		TC

TABLE 1—*Continued*

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM*	Calc. FOM*	Reference
ATC/CTC	I269L	No predictable effect of substitution.	4.43		(50)
ATC/AAC	I269N	No room for an Asn.			TC
AGA/AAA	R270K	Substitution to a Lys or a Ser may break salt-	3.78		(71)
AGA/AGT	R270S	bridge to Asp282 and a H-bond to the carbonyl- oxygen of Pro275.			(72,18)
CAT/TAT	H271Y	Active site defining H-bonds made to Asp161 and Ser273 will be broken if substituted to a Tyr.	5.85		(36)
TCC/TTC	S273F	Active site defining H-bond to His271.	3.75		(73)
CCC/TCC	P275S	Substitution of a Pro in the active site into a Ser may change the shape of the active site.	3.45		TC
ATG/ATT	M276I	On surface of CD.	7.63		(45)
ATG/GTG	M276V	Exposing hydrophobic residues to surface.			(74)
TAT/GAT	Y277D	On surface. No predictable effect of	6.61	1.77	(75)
TAT/TGT	Y277C	either substitution.			(71)
ACC/AAC	T278N	At entrance to active site. Any			(39)
ACC/ATC	T278I	substitution would break a H-bond to	2.09		(76)
ACC/GCC	T278A	Glu280 and affect shape of active site.			(71)
GAA/AAA	E280K	Forms a saltbridge to Arg158 plus a H-bond to His146. Substitution would change the channel width of active site.	7.11	14.14	(19,20)
CCg/CTg	P281L	Helps define the shape of the active site in immediate vicinity to the iron. Substitution would change structure of this area.	3.62	22.98	(22,23)
GAC/AAC	D282N	H-bonds to Arg270. Helps to define shape of active site.	4.83		(39)
ATC/TTC	I283F	No room for larger sidechain than an	4.28		(77)
ATC/AAC	I283N	Ile.			(50)
CAT/TAT	H285Y	Fe-ligand!!	5.21		TC
TTG/TTC	L288F	Possibly not enough room for a Phe.	3.04		TC
CGC/CAC	R297H	Hydrogen bonds to Arg71 and Glu422 in other subunit of dimer. Substitution to His would break	15.58		TC
CGC/TGC	R297C	hydrogen bonds.			(28)
TTT/TGT	F299C	π-stacks onto Phe392, which again π -stacks onto Tyr343. At bottom of active site.	3.57	12.37	(78,20)
GCC/GTC	A300V	Close to Thr 238. Not enough room for			(66)
GCC/TCC	A300S	larger sidechains of Val and Ser.	4.00	5.30	(79)
TCC/CCC	S303P	Located close to (behind) His 290. A substitution	3.42		TC
TCC/GCC	S303A	to a Pro will disturb active site structure.			TC
CAG/CGG	Q304R	Substitution to an Arg would break a hydrogen bond to Arg261 and a hydrogen bond to Tyr417 in other subunit of dimer.			TC
ATT/GTT	I306V	No predictable effect of substitution.	3.80	7.07	(80)
CTT/GTT	L308V	No predictable effect of substitution.	4.11		TC
GCC/GAC	A309D	At end of Cα8. Substitution to a larger			(81)
GCC/GTC	A309V	residue would push helix Cα12 away.	4.00	1.77	(39)
TCT/TTT	S310F	A hydrogen bond to the carbonyl-oxygen of Gly307 helps to keep Cα8 and Cα9 together.	2.93		TC
CTG/CCG	L311P	Substitution to a Pro may break hydrogen bond to Arg 408 important for positioning TD helix.	3.83	1.77	(82)
CCT/CAT	P314H	At start of helix Cα9. A substitution to a His would make it more difficult to start helix.	3.62		(61)
AAG/AAC	K320N	On surface. No contacts. No predictable effect of substitution.	2.20		(25)
GCC/ACC	A322T	Close to Ser251. Controls			(79)
GCC/GGC	A322G	separation of helix Cα6 & Cα9.	6.43	10.61	(35)
TTT/TTG	F327L	Stacks between Phe387and Trp326 plus Tyr377.	3.03		TC
ACT/GCT	T328A	Substitution to a Leu would disrupt π-stacking . Hydrogen bonds to Tyr325 and Ile324 as well as	2.11		TC

TABLE 1—*Continued*

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM*	Calc. FOM*	Reference
		Val 329. These are all important for keeping the overall shape of active site intact.			
TTT/CTT	F331L	Stacks onto Trp326. Contributes to the wall of the	2.99	1.77	(24)
TTT/TGT	F331C	active site. Any substitution would abolish π-stacking interactions and disrupt packing here.			(25)
CTC/TTC	L333F	Can easily accomodate a Phe with slight movement of end of Cα11 helix.	3.24		TC
TGC/TCC	C334S	Forms a disulfide bridge to Cys203.	11.69		TC
GGA/GTA	G337V	Between Cβ3 and Cβ4. A Gly may be neccesary to form this hairpin loop.	7.85		(36)
GAC/TAC	D338Y	Also between Cβ3 and Cβ4. No predictable effect of substitution.	4.12		(81)
AAG/ACG	K341T	In Cβ4. No contacts. No predictable effect of	2.71		(36)
AAG/AGG	K341R	substitution.			TC
GCA/ACA	A342T	At end of Cβ4. Ala may be neccesary to have correct loop formation. Substitution to a Pro	3.71		(39)
GCA/CCA	A342P	changes secondary structure close to active site.			TC
TAT/TGT	Y343C	π-stacks onto Phe392 which π-stacks onto Phe299.	11.01		(39)
TAT/TTT	Y343F	No predictable effect of substitution to a Phe, but a Cys may destabilize the stacking interactions.			TC
GGT/AGT	G344S	No room for a Ser. In loop before Cα10. May be neccesary to start helix close to active site Fe.	7.12		TC
GCT/ACT	A345T	Ala345 is right behind His285, His290 and Glu330 (Fe-ligands). There is no room for a	4.17		(49)
GCT/TCT	A345S	substitution to a larger residue. Will interfere with Fe-binding.			TC
GGG/CGG	G346R	Same as above.	6.89		TC
CTC/TTC	L347F	No room for a substitution to a Phe. Close to active site	3.24		(25)
CTG/GTG	L348V	No predictable effect of substitution.	3.88	3.54	(78)
TCA/CCA	S349P	H-bonds to His285, which keeps	7.46	3.54	(83,26)
TCA/TTA	S349L	the His in place to bind Fe.		1.77	(27)
TCC/ACC	S350T	At start of Cα11 helix. Hydrogen bonds through a water molecule to Glu353 at active site edge.	3.75		TC
GGT/CGT	G352R	Solvent exposed. At start of Cα11. May be neccesary to get helix started properly.	7.12		TC
TAC/CAC	Y356H	In Cα11. No contacts made. No predictable effect of substitution.		1.77	TC
CCA/ACA	P362T	In an area with no secondary structure. Both Pro 362 and Pro366 may be important for mediating	2.95		(38)
CCC/CAC	P366H	dimer contacts between Glu368 in A and B subunits.	3.70		(39)
AAG/AGG	K371R	On surface. No predictable effect of substitution.	2.74		(52)
ACA/TCA	T372S	No predictable effect of substitution.	2.94		(61)
TAC/TGC	Y377C	Stacks onto Trp326. H-bonds to Ser23 in RD. Important for regulating access to active site.	7.20		TC
ACG/ATG	T380M	H-bonds to Ser378.	9.69	8.84	(39)
TAT/TGT	Y386C	H-bonds to Glu353 in Cα11.	6.41		(64)
TAC/CAC	Y387H	H-bonds to Gln375.	8.24		(52)
GTG/ATG	V388M	On surface. No predictable effect of	7.01	21.21	(84)
GTG/CTG	V388L	substitutions.			TC
GAG/GGG	E390G	No contacts. On surface. No predictable effect of substitution.	4.64	3.54	TC
GAT/CAT	D394H	Substitutions result in disruption of	5.76		(25)
GAT/GCT	D394A	H-bonds to Lys398 in Cα12.			(39)
GCC/CCC	A395P	A substitution to a Pro will disturb helix Cα12 structure. No predictable	4.00	1.77	(39)
GCC/GGC	A395G	effects of substitution into a Gly.			TC
AGG/CGG	R400R[s]	Silent mutation. Can affect splicing.	4.37		TC

TABLE 1—*Continued*

cDNA mutation	PheOH mutation	Structural contacts/comments	Pred. FOM[%]	Calc. FOM[&]	Reference
GCT/GTT	A403V	At end of helix Cα12. May be neccesary for starting loop before Tβ1.	4.17	7.07	(85)
ATA/ACA	I406T	No predictable effects of substitution.	4.04		TC
CCT/TCT	R407S	Arg may be important for positioning TD helix.	3.62	1.77	(53)
CGG/CAG	R408Q	In hinge loop which connects TD arm to CD. H-bonds to Leu 311 and Leu308 carbonyl-oxygens.	21.25	14.14	(35)
CGG/TGG	R408W	Substitution to any other sidechain disrupts important hydrogen bonds.		135.21	(34,35)
CGC/AGC	R413S	In middle of Tβ1. No contacts.	15.58		TC
CGC/TGC	R413C	Substitution to a Pro disturb			TC
CGC/CCC	R413P	packing in tetramerization arm.			(86)
TAC/TGC	Y414C	H-bonds through 2 water molecules to Phe260 and Gly239 carbonyl oxygens. Important for keeping TD close to CD.	8.77	82.19	(23)
GAC/AAC	D415N	H-bonds to Thr418. Stabilizes loop between Tβ1 and Tβ2.	6.26	15.91	(80)
ACC/CCC	T418P	Same as above. Substitution to Pro disturb loop structure and positioning of tetramerization helix.	2.72		(49)
GAG/AAG	E422K	H-bonds to Arg297 in other subunit and RD.	5.48		TC
CTT/CCT	L430P	At start of Tα1. Substitution to Pro disturbs helix and tetramer formation.	3.90		TC
GCC/GAC	A447D	Substitution to Asp will disrupt coiled-coil interactions in TD.	4.00		(25)

Note. E; exon, I; intron; X, nonsense mutation, inserts premature stop codon; ND, not detectable; Hkc, human kidney cells; NS, not stated; RD, regulatory domain; CD, catalytic domain; TD, tetramerization domain; Euk., eukaryotic cells (COS cells); Non-PKU HPA definition, blood plasma L-phe levels between 120 and 1000 μM; Classical PKU, blood plasma L-phe levels above 1000 μM; Variant PKU, blood plasma L-phe levels fits neither classical PKU nor non-PKU HPA; TC, reference not published in any journals. Submitted to the Phenylalanine Hydroxylase Mutation Analysis Consortium Database (1998).

[%] Predicted frequency of mutation (FOM) based on the program MUTPRED (15)—weighted data.

[&] Calculated frequency of mutation (FOM) based on total amount of known mutation and their frequency in 496 patients currently phenotyped. Data taken from Phenylalanine Hydroxylase Mutation Analysis Consortium Database (1998) information on PKU genotype versus phenotype.

Calculated FOM: α = {No. of times a mutation occurs on one allele/No. patients for which data exist (= 495)} · 100%.

Calculated FOM (normalized to predicted FOM) = $\alpha \cdot \{\Sigma$ (predicted FOM (weighted))/Σ (α_{aver} (for each residue)}. α_{aver} = averaged calculated FOM for residues that have more than one mutations known.

ing both kinetic and stability properties of the enzyme; (ii) structurally stable mutations with altered kinetic properties; and (iii) mutations with normal kinetics, but reduced stability *in vitro* and *in vivo* (14). Future continuation of these expression analyses, in parallel to the known three-dimensional structure of PheOH, is important for a detailed characterization of the catalytic and regulatory properties of the disease-causing mutants in the *PAH* gene.

Single-Point Mutations

There are currently 234 missense point mutations known for the *PAH* gene (PAH Mutation Analysis Consortium Newsletter, Spring 1998), in addition to 22 nonsense mutations and 11 silent mutations. Most of the single-point mutations map onto the region consisting of exon 5 (residue 148) to exon 12 (residue 438). Forty-nine of the total number of mutations are located in the regulatory domain (residues 1–142), 209 mutations are located in the catalytic domain (residues 143–410), and 10 are located to the tetramerization domain (residues 411–452). A summary of the known genotype/phenotype information about these mutations is shown in Table 1 (illustrated in Fig. 3), together with a structural interpretation of the predicted effect of the selected missense mutations on the enzyme function.

In an attempt to relate the variable prevalence of human genetic disease (PKU) to the mutability inherent in the nucleotide sequence of the *PAH* gene, a

mutability profile of the PAH cDNA sequence was performed using the program MUTPRED (15). The results are presented in Table 1 for the residues in PheOH that are found to result in single-point mutations in the PKU database. The silent mutations (G10G, S137S, C203C, Q232Q, V245V, Q304Q, T323T, L385L, K398K, V399V and Y414Y) are not shown in the table since these do not introduce changes in the protein structure. The nonsense mutations (Q20X, Y77X, R111X, W120X, Y166X, Q172X, R176X, W187X, Y204X, Y206X, Y216X, Q232X, R243X, R261X, G272X, S295X, W326X, Q336X, K341X, Q355X, Y356X, S359X) are not shown in the table since these mutations lead to a premature step in translation. The MUTPRED program generates unweighted or weighted profiles for which the relative mutabilities of CpG dinucleotides are multiplied by the likelihood of clinical detection depending on chemical difference in the translated amino acid residues (Fig. 3B). It is based on a higher mutation frequency in CpG dinucleotides than in other dinucleotides, thus causing human disease. To compare this theoretical frequency of mutation to the actual detected frequency of mutation in the *PAH* gene, an analysis of the known mutations in the PKU database was done for the single-point mutations known (Fig. 3B). These data are also presented in Table 1 as calculated frequency of mutations in the *PAH* gene.

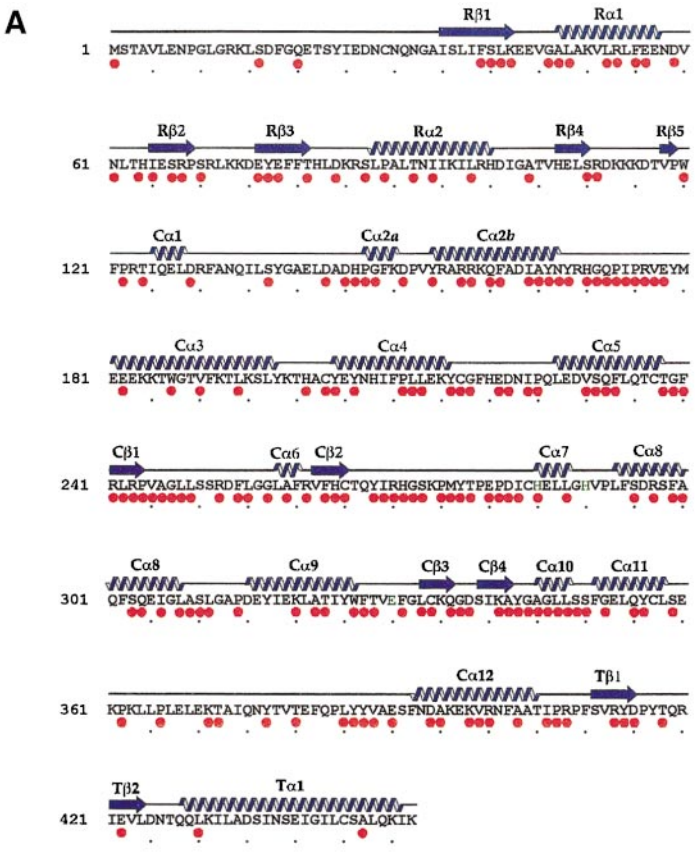
STRUCTURAL INTERPRETATION OF PREVALENT MUTATIONS

Structurally, the mutations causing PKU/HPA can be shown to affect residues in five categories: (a) active-site residues, (b) structural residues, (c) interdomain interactions in a monomer, (d) residues interacting with the N-terminal autoregulatory sequence, and (e) residues in the dimer or tetramer interfaces. Thus, the most prevalent single-point mutations in the *PAH* gene will be discussed according to these five groups.

Active-Site Mutations

Thirty-one missense point mutations can be found in the residues lining the active site. Seven of these can be found in the putative pterin-binding motif, consisting of residues 264–290 (16). Five of the active-site mutations have been expressed *in vitro*, some even in multiple *in vitro* systems. The first residue in the catalytic domain (Asp143) is reported to be a PKU mutant associated with the severe PKU phenotype (17) in patients heterozygous for the termination mutant G272X and the D143G mutation. This residue is located to the surface at the entrance to the active site and may be involved in controlling access to the active site for substrate/cofactor. *In vitro* expression of the mutant in three different systems (*Escherichia coli*, human kidney cells, and a transcription-translation system (TnT)) resulted in a relatively high residual enzymatic activity (52, 33, and 102%, respectively) as compared to the wild-type PheOH. Also, when assayed at lower concentrations of L-Phe and BH_4, the residual activities were compatible with severely reduced hydroxylation of L-Phe, thus suggesting that this residue indeed plays a role in controlling substrate access to the active site. Another one of the active-site residues which may be involved in the binding of the BH_4 cofactor in PheOH; Phe254 is found to π-stack on the pterin ring in the structure of TyrOH complexed with 7,8-dihydrobiopterin (12). In the structure of PheOH, this residue is located 5.9 Å away from the active-site iron, and a substitution into a isoleucine as in the PKU mutant F254I would most likely interfere with proper binding of pterin and possibly also the phenylalanine substrate. A frequent mutation among black Americans is the L255S mutation, which causes severe PKU (18). In this mutation a nonpolar group is substituted with a smaller polar side chain in a hydrophobic environment and this is expected to result in significant structural perturbation. Leu255 seems to be important in controlling the separation of helix Cα6 and Cα9. Expression of the two mutations associated with this residue, L255S and L255V, in three different *in vitro* expression systems, showed that, as expected, the L255S mutation has a far larger effect on the residual enzyme activity (only 1% residual activity detected in the TnT expression system versus wild-type) than the more conserved L255V mutation, which revealed 11% residual activity (TnT expression system), 44% residual activity (*E. coli*), and 13% residual activity (COS cells), versus wild-type PheOH (18). As can be expected, based on the more conserved mutation, the L255V mutation is known to cause mild PKU.

The active-site residue Thr278 has three different mutations associated with it. Thr278 is located at the entrance to the active site and hydrogen bonds to Glu280 (Fig. 2). The mutations T278I and T278A are substitutions of a polar amino acid into hydrophobic amino acids and will result in this important hydrogen bond being lost and also perturb the structure at the active-site entrance. The more electrostatically conserved mutation T278N would putatively also


A
Rβ1
Rα1
1 MSTAVLENPGLGRKLSDFGQETSYIEDNCNQNGAISLIFSLKEEVGALAKVLRLFEENDV
Rβ2
Rβ3
Rα2
Rβ4
Rβ5
61 NLTHIESRPSRLKKDEYEFFTHLDKRSLPALTNIIKILRHDIGATVHELSRDKKKDTVPW
Cα1
Cα2a
Cα2b
121 FPRTIQELDRFANQILSYGAELDADHPGFKDPVYRARRKQFADIAYNYRHGQPIPRVEYM
Cα3
Cα4
Cα5
181 EEEKKTWGTVFKTLKSLYKTHACYEYNHIFPLLEKYCGFHEDNIPQLEDVSQFLQTCTGF
Cβ1
Cα6
Cβ2
Cα7
Cα8
241 RLRPVAGLLSSRDFLGGLAFRVFHCTQYIRHGSKPMYTPEPDICHELLGHVPLFSDRSFA
Cα8
Cα9
Cβ3
Cβ4
Cα10
Cα11
301 QFSQEIGLASLGAPDEYIEKLATIYWFTVEFGLCKQGDSIKAYGAGLLSSFGELQYCLSE
Cα12
Tβ1
361 KPKLLPLELEKTAIQNYTVTEFQPLYYVAESFNDAKEKVRNFAATIPRPFSVRYDPYTQR
Tβ2
Tα1
421 IEVLDNTQQLKILADSINSEIGILCSALQKIK
B
R408
Y414
R243
R261
I65
G46
R111
R169
R176
W120
W187
L48
R158
R155
R252
P281
R241
W326
R408
R413
V388
A104
E280
R297
F299
D415
0
20
40
60
80
100
120
140
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450

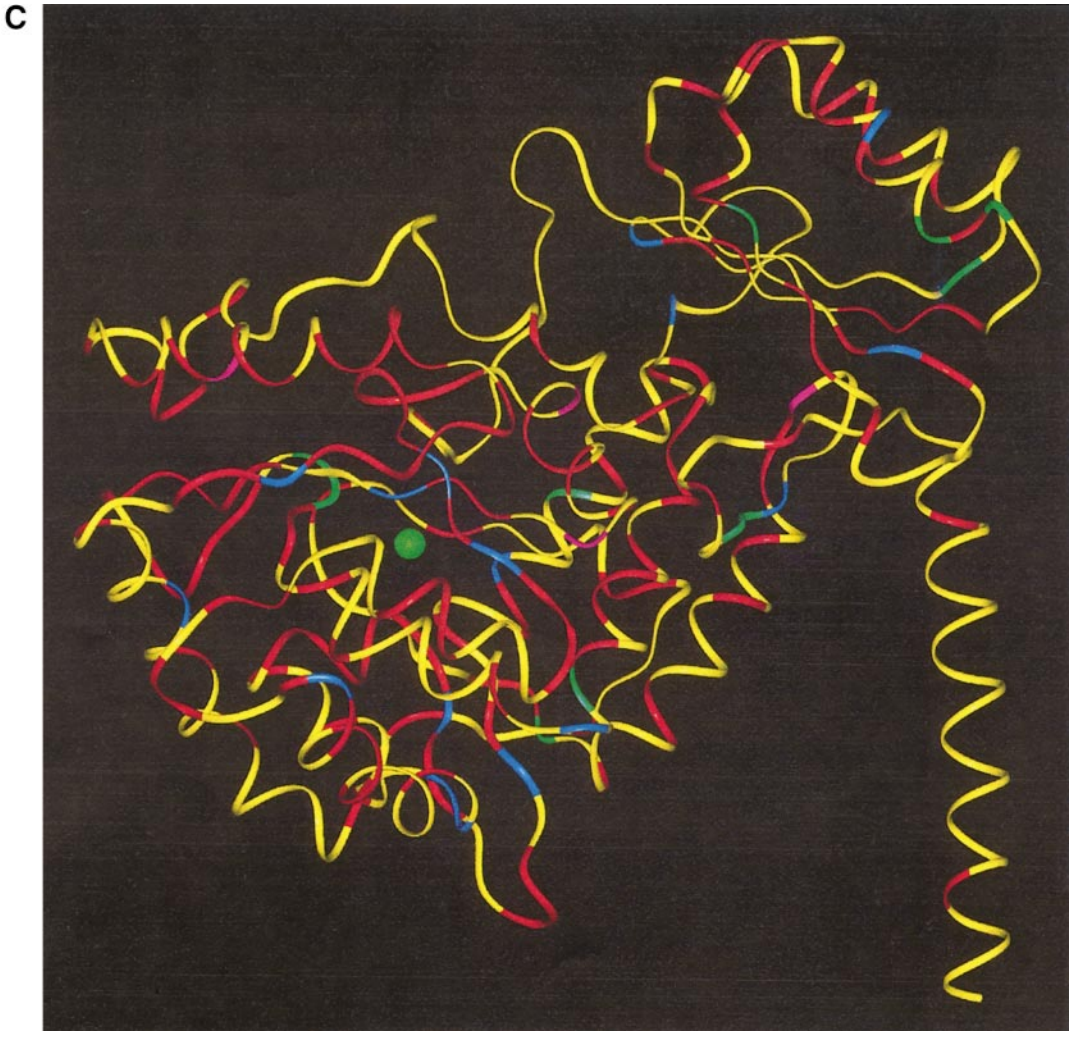


FIG. 3. (A) Scheme showing the secondary structure assignment of the human PheOH sequence (SWISS-PROT P00439). The secondary structure was assigned using DSSP (43) on the composite PheOH model. The residues that have PKU mutations associated with them are marked with red dots. (B) Predicted mutability profile for the human PAH gene calculated by the program MUTPRED (15). Also shown for comparison (in red circles) is the calculated frequency of mutation based on the known frequency of mutations occurring in 496 alleles causing HPA and PKU (genotype/phenotype data compiled by E. Kayaalp (PAH Mutation Analysis Consortium Newsletter, Spring 1998, and Ref. 44)). The data are listed in Table 1 for all known single-point mutations causing HPA/PKU. (C) Cα-trace of one monomer of the full-length model of phenylalanine hydroxylase. The trace is colored yellow for residues not associated with any PKU mutations, and the active-site iron is shown for reference as a green sphere. The region colored red consists of residues that have PKU mutations associated with it. The regions colored blue have residues that show a predicted frequency of mutation (as calculated by MUTPRED) (Table 1) higher than 10, and the regions colored green have residues that show a calculated frequency of mutation (calculated from data compiled by E. Kayaalp and Ref. 44) higher than 10 (Table 1). The four residues colored magenta are the residues with both the calculated and the predicted frequency of mutation being high (Arg158, Arg252, Arg261, and Arg408).

result in a change of hydrogen-bonding pattern, and thus have the same perturbing effect as that of the two other mutations, if only slightly smaller.

Another well-studied mutation located to the active site is the E280K mutation. This residue is also found within the so-called pterin-binding motif (16), and the mutation itself causes mild to severe PKU (19). In the PheOH structures, Glu280 is seen to hydrogen-bond to His146 and to form an important salt bridge to Arg158 (Fig. 2). Also, there are only two free, charged groups in the active-site crevice of PheOH (both glutamic acids) and the substitution of one of them (Glu280) by a lysine represents a dramatic change in the electrostatic potential in the active site. The expression of this mutation in human PheOH in *E. coli* resulted in an enzyme with

only ~1% of the specific activity of wild-type PheOH (20) as assayed with both the natural cofactor BH_4 and the synthetic cofactor 6-methyltetrahydropterin (6-MPH_4).

One of the most frequent mutations in southeastern Europe is the P281L mutation (21), causing phenotypes ranging from HPA to severe PKU. This proline helps to define the shape of the active site very close to the iron in the PheOH structure (Fig. 2). It is therefore almost certain that a substitution to a less rigid leucine will change the conformation of the active site by removing the conformational constraints imposed by the proline. *In vitro* expression of this mutant in *E. coli* resulted in complete loss of PheOH activity (22), and expression in COS cells resulted in <1% enzyme activity (23), as compared to wild-type PheOH.

Three residues that are in the active site are Phe331, Ala345, and Gly346. Phe331 π-stacks onto Trp326, and both Ala345 and Gly346 are located right behind the triad of residues that are responsible for binding the iron (His285, His290, and Glu330) (Fig. 2). Two PKU mutations have been reported in the Phe331 residue: F331C and F331L (24,25). Both mutations will abolish important π-stacking interactions that stabilize the wall of the active site. Ala345 has two reported PKU mutations, A345S and A345T, and Gly346 has one, G346R. All three mutations increase the size of the side chain which would putatively result in the iron ligands being shifted further out into the active-site crevice, and thus destabilize the entire active-site structure.

One more residue which seems very important in holding the iron in place is Ser349, which hydrogen-bonds to His285, one of the iron ligands. Two PKU mutations have been reported for this residue, one of them resulting in classical PKU (S349P) (26). This substitution of a proline so close to the catalytic iron will change the entire shape of the active site. The mutation has been expressed in both *E. coli* and COS cells and has been found to have only <0.2 to <1% residual activity and no detectable levels of activity, respectively, as compared to the wild-type PheOH. The second mutation in this residue (S349L) (27) is also reported to give similar *in vitro* expression in *E. coli* and COS cells of <0.2% or no detectable levels of PheOH activity, respectively. Both mutations disrupt the very important hydrogen bond to His285, and thus have severe influences on the catalytic activity of PheOH.

Structural Residues

The G46S mutation in the regulatory domain is associated with classical PKU and is one of the most frequent mutations found in the PKU database (Table 1). The cDNA site for residue 46 has also been predicted to be a site of mutation by the program MUTPRED. Gly46 is located on the surface of the regulatory domain in a loop just before helix Rα1. The substitution with a serine seems to involve the serine side-chain hydrogen bonding to one or more putative residues in close proximity to Gly46 in the PheOH structure, and thus results in distortions of the secondary structure in the regulatory domain. This could result in the formation of inactive aggregates as reportedly found by Eiken *et al.* (28) upon expressing the mutant in three different systems (*E. coli*, human kidney cells, and TnT). Another frequent mutation found in the regulatory domain is the I65T mutation. This mutation is associated with non-PKU HPA to variant PKU and has been expressed in COS cells to find a residual activity of approximately 26% of wild-type PheOH (29). Ile65 is located in the hydrophobic core of the regulatory domain. A substitution with a polar threonine (or asparagine as in the I65N mutant) would result in significant structural perturbation. A more conserved substitution, like in the I65A substitution, would probably not cause as large a perturbation as the two others. Another mutation of relatively high frequency in patients with PKU is the A104D mutation. This mutation is associated with variant PKU, and the Ala104 residue is located in the loop between Rα2 and Rβ4 in the regulatory domain. A substitution by a larger and charged residue may destabilize the loop structure. When expressed *in vitro* in three different systems, the increased aggregation of the resulting enzyme was observed, and residual activity as compared to the wild type varied from 26% (human kidney cells) to no detectable levels (*E. coli*) (30).

Arg158 is the site of predicted frequent mutation as determined by MUTPRED based on the cDNA sequence of PheOH (15). The R158Q mutation is associated with classical PKU and is also a frequent mutation in patients with PKU (Table 1). Arg158 forms a salt bridge to Glu280 and one hydrogen bond to Tyr268 (Fig. 4A). Both of these interactions are important for conserving the shape of the active site, and substitution with a glutamine, or the larger aromatic residue, tryptophan (R158W mutation), will necessarily distort the active-site structure. *In vitro* expression of the R158Q mutant was found to

result in approximately 10% of normal levels of PheOH activity (31).

Cys203 is found to be involved in a disulfide bond with Cys334 in the structure of rat PheOH including the regulatory domain (11). This is not observed in any of the structures of human PheOH (9,10), but that may be a product of the expression system (insect cells versus *E. coli*). Both cysteines involved in the disulfide bond are associated with PKU mutations and are predicted to be sites of disease-causing mutations (by MUTPRED) (15). Another predicted site of disease-causing mutation is in Arg243. This residue is located at the end of Cβ1 and forms a salt bridge to Asp129 in Cα1 helix. *In vitro* expression of the two PKU mutants associated with this residue (R243Q and a nonsense mutant R243X) in COS cells resulted in $<$10% residual activity (R243Q) (32) and $<$1% residual activity (R243X) (19). Expression of the R243X mutant in *E. coli* resulted in activity levels similar to those in COS cells ($<$0.2%) (20). Arg252 forms a salt bridge to Asp315 and hydrogen-bonds to the carbonyl oxygen of Ala313 as well as the side chain of Asp27 in the autoregulatory sequence (Fig. 4B). There are three PKU mutations associated with this residue: R252Q, R252G, and R252W. The latter of the three mutations is reported to result in classical PKU, and *in vitro* expression of the R252G and R252W mutations in *E. coli,* TnT, and COS cells results in $<$1% residual activity as compared to that in the wild-type PheOH (18,20,31). The R252Q mutation, however, results in somewhat larger residual activity (16 and 3%) when expressed in *E. coli* and TnT systems (18). Thus it is quite apparent that any substitution in this residue will result in disruption of stabilizing interactions in the catalytic domain.

The residue Ala259 is buried in a hydrophobic pocket about 4 Å away from Leu311 and Leu308, which are hydrogen-bonded to Arg408. These residues are important for holding the tetramerization arm close to the catalytic domain, and thus a substitution of Ala259 into a larger or polar residue, as in the PKU mutations A259T and A259V, cannot be accommodated without disturbing the surrounding environment (including Leu311 and Leu308). Both mutations have been expressed *in vitro* in *E. coli* and result in 0.1–0.8% residual activity as compared to that in wild-type PheOH (18,20).

One of the most frequent mutations based on the data in the PKU database (also predicted by MUTPRED) (15) is the R261Q PKU mutation. This mutation is reported to result in phenotypes varying from variant PKU to the more severe form, classical PKU. Arg261 is located in a loop between helix Cα6 and Cβ2 in the catalytic domain, where it hydrogen-bonds to Gln304 and Thr238 (Fig. 4C). This helps to stabilize the structure of the active site, and substituting the arginine to a proline (R261P) would necessarily destabilize the active-site structure. *In vitro* expression of the R261Q mutant in *E. coli* (20) and COS cells (19) results in comparable levels (32 and 30%, respectively) of enzymatic activity. Thus, the destabilization of the active-site structure is not enough to completely abolish PheOH activity.

One interesting mutant within the putative pterin-binding region in the catalytic domain is the nonsense mutant G272X. Expression analysis of this mutant results in no detectable PheOH activity in either *E. coli* or COS cells (17,18,33). Glu272 is located in a loop between Cβ2 and Cα7 just before the active-site histidines that bind the catalytic iron, and the mutation would putatively result in a truncated form of PheOH that has none of the residues that are responsible for binding iron, thus no catalytic activity should be observed.

Interdomain Interactions in a Monomer

The most important residues at the interface between the catalytic domain and the tetramerization domain are residues Leu311, Leu308, and Arg408. Arg408, which is in the loop between Cα12 and Tβ1, is buried and forms hydrogen bonds to the carbonyl oxygens of Leu308 (at the end of Cα8) and Leu311 (in the loop between Cα8 and Cα9) (Fig. 5A). This residue is predicted to have a high frequency of mutation as predicted by MUTPRED (15) (Table 1), but is also the site of the single most frequent mutation in the PKU database (R408W mutation). The mutation results in a severe metabolic PKU phenotype and is reported to give low ($<$1 to $<$2.7%) residual activity when expressed in COS cells (34,35).

A second frequent mutation in this residue is the R408Q mutation, which is associated with a less severe metabolic non-PKU HPA phenotype and reportedly gives 55% residual PheOH activity, as compared to that of the wild type, if expressed in COS cells (35). The substitution of Arg408 into the larger and bulkier tryptophan would alter the hydrogen-bonding network that holds the tetramerization domain together with the catalytic domain, and through steric hindrance interfere with the correct positioning of the β-ribbon (Tβ1 and Tβ2).

The mutation Y414C has been found to be a relatively frequent mutation and result in non-PKU

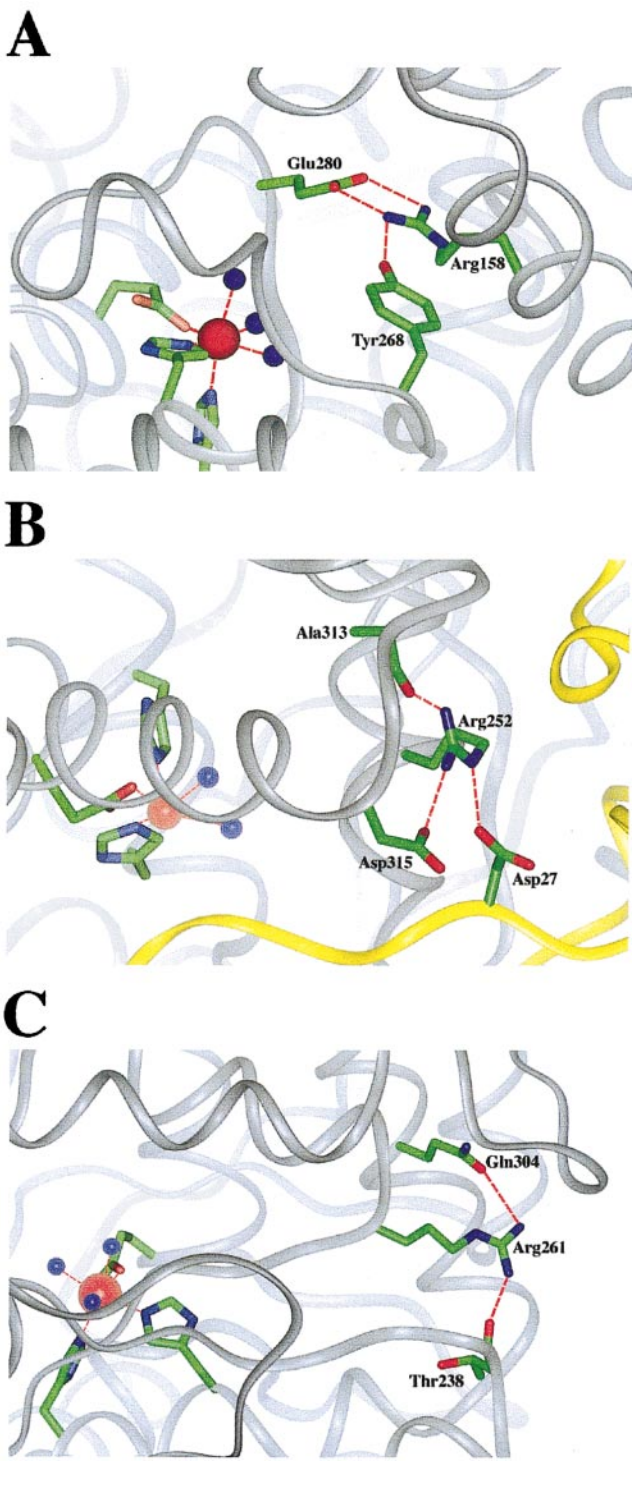


FIG. 4. Various active-site residues that have PKU mutations associated with them, including their hydrogen-bonding partners. The active-site iron and its ligands are also shown for reference (A) Arg158, (B) Arg252, and (C) Arg261.

HPA to variant PKU phenotypes. *In vitro* expression in COS cells results in an enzyme which retains 28–50% of normal activity (23). The mild effect of this mutation can be observed also at a structural level; a substitution from a tyrosine to a cysteine in this position should not have drastic consequences on the interaction with the neighboring residues.

Residues Interacting with the N-Terminal Autoregulatory Sequence

One of the few residues in the catalytic domain that seems to be in contact with the N-terminal autoregulatory sequence is Tyr377 (11). This residue is found to hydrogen-bond to Ser23 in the regulatory domain and also stacks on top of Trp326 (Fig. 5B). It should be important for regulating access to the active site and therefore also regulates the substrate and cofactor access to the catalytic iron. In the PKU mutant Y377C, the important hydrogen bond would be disrupted, thus exposing the active site and destroying the effects of the regulatory domain upon substrate binding. As mentioned previously, Asp27 hydrogen-bonds to Nϵ of Arg252, and in the PKU mutations associated with Arg252 (R252Q, R252G, and R252W) the disruption of this hydrogen bond will most likely also change access to the active site for substrate/cofactor.

Residues at the Dimer or Tetramer Interfaces

Several residues that have PKU mutations associated with them can be found at the dimer interface in all structures of PheOH solved so far (9–11). These are residues Ser67, Glu76, Arg297, Gln304, and Glu422. Ser67 is located in β-strand Rβ2, and its main chain amide hydrogen-bonds to the hydroxyl group of Tyr216 in the other monomers that make up the dimer. The PKU mutation S67P would necessarily distort the structure in this region, imposing more rigid torsion angles than those of the native serine (36). Glu76 is located on the surface of the regulatory domain, in β-strand Rβ3, where the main-chain carbonyl oxygen of Glu76 in one monomer is hydrogen-bonded to Nϵ2 of His208 in the second monomer, as well as Asn73, just prior to Glu76 in sequence (Fig. 6A). This contact seems to be important for forming the dimer that makes up the "dimer of dimers" seen in the tetrameric structure of PheOH (10). Glu76 has two PKU mutations associated with it, E76A and E76G (Table 1). Both of these mutations are substitutions of a charged amino acid into surface-exposed hydrophobic residues, and this could result in perturbations of the

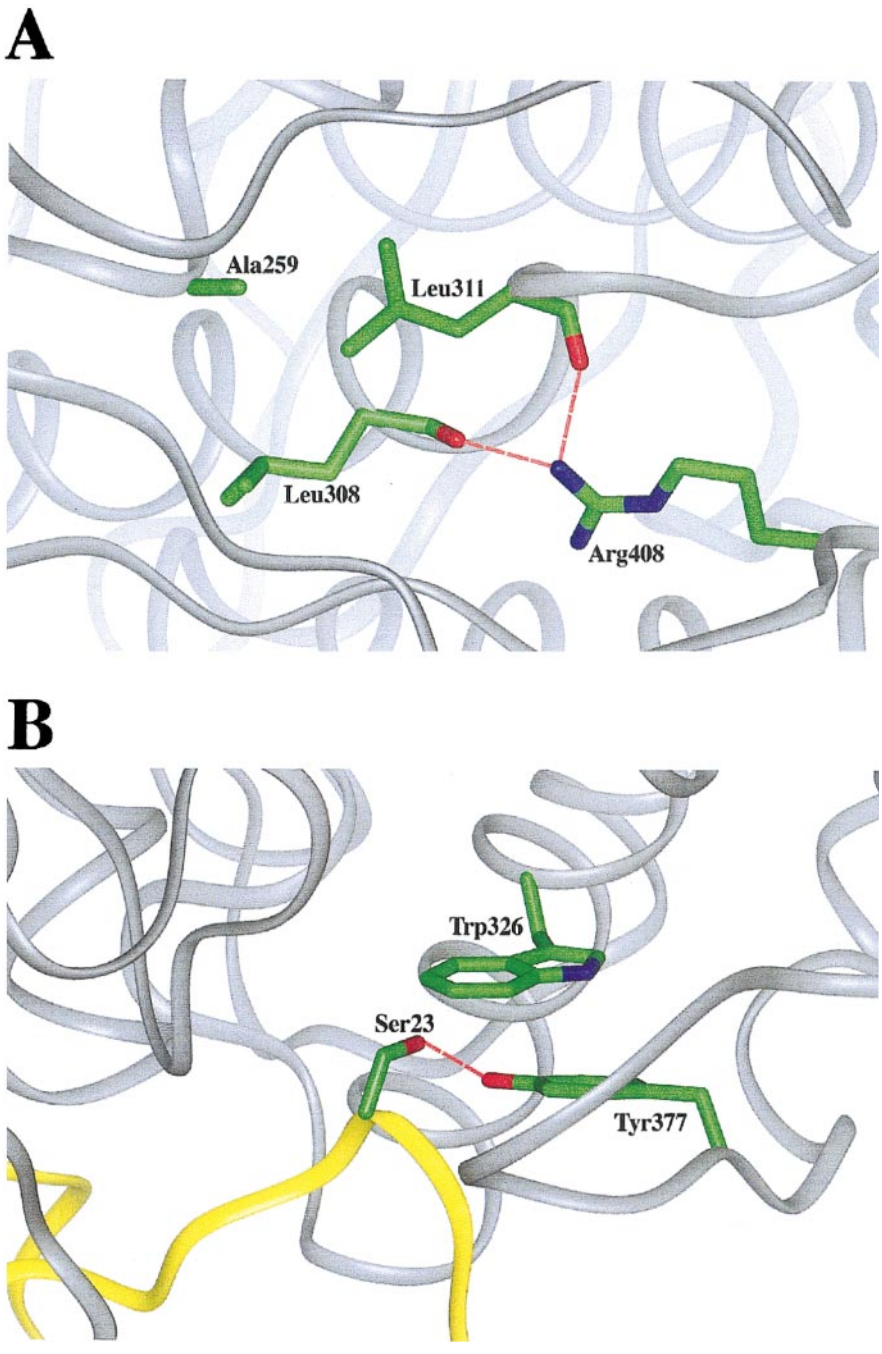


FIG. 5. (A) Arg408 is the site of the most frequent PKU mutation (R408W). It is hydrogen-bonded to Leu311 and Leu308 in the same PheOH monomer. This residue seems to be important for holding the catalytic domain together with the tetramerization domain. (B) The PKU-mutated residue Tyr377 is associated with the N-terminal regulatory sequence. It hydrogen-bonds to Ser23 and π-stacks onto Trp326. The N-terminal regulatory domain Cα-atom trace is shown in yellow ribbon, whereas the catalytic domain Cα-atom trace is shown in gray ribbon.

secondary structure in the regulatory domain. Arg297 is located in α-helix Cα8 and is predicted (by MUTPRED) (15) to be a site of frequent mutation. The side chain of Arg297 is hydrogen-bonded to the main-chain carbonyl oxygen of Arg71 as well as the side-chain oxygen of Glu422 in the other dimeric monomer (Fig. 6B). There are two PKU mutations listed for Arg297 in the PKU database, R297H and R297C (28). Both substitutions will result in the disruption of the above-mentioned dimer-stabilizing

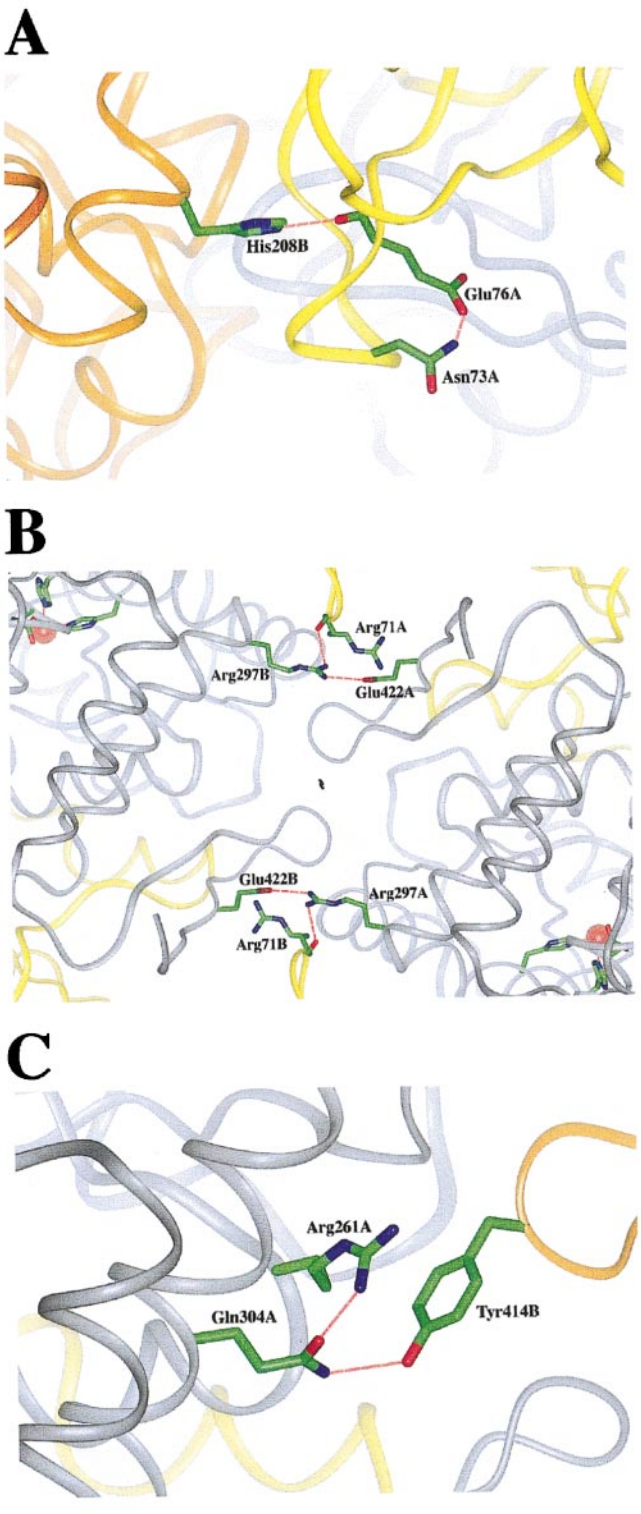


FIG. 6. Various PKU-associated residues at the tetramer/dimer interfaces. (A) Glu76 with its hydrogen-bonding partners Asn73 and His208 (of monomer B). The Cα-trace of the catalytic domain and the regulatory domain in one monomer are shown in gray and yellow, respectively, whereas the other monomer in the dimer is shown in orange to ease comparisons. (B) Arg297 with its hydrogen-bonding partners Glu422 and Arg71 (of monomer B). The presence of a twofold symmetry axis perpendicular to the plane of the figure is shown in the middle of the picture. Both Cα-traces of the catalytic domains in each monomer are shown in gray, and the Cα-trace of the regulatory domain is shown in yellow. The catalytic 2-His-1-Glu triad of coordinating residues to the iron is shown for structural reference. (C) Gln304 with its hydrogen-bonding partners Arg261 and Tyr414 (of monomer B). The same color scheme as in (A) is employed.

hydrogen bonds and would possibly result in the formation of unstable PheOH enzyme. The last residue associated with PKU mutations *and* involved in dimer/tetramer interactions is Gln304. This residue is located in α-helix Cα8 and forms hydrogen bonds to the hydroxyl group of Tyr414 (in monomer B) (Fig. 6C), as well as to Arg261. Gln304 has two PKU mutations associated with it, one silent mutation (Q304Q) (37) and Q304R. The substitution of the polar glutamine to a charged arginine could disrupt both hydrogen bonds and destabilize the PheOH enzyme formed.

Lastly, both Pro362 and Pro366 seem to be important for positioning Glu368 correctly in order to mediate dimer contacts between the same Glu368 in two different subunits. Thus the PKU mutations P362T (38) and P366H (39) will both destabilize these parts of the structure and prevent proper dimers from being formed.

Deletion and Insertion Mutants

Most of the non-single-point mutations that may affect the structure of PheOH are listed in Table 2. These include deletion, insertion, and one prevalent intron mutation which affect the structure of PheOH. The frameshift mutations have been excluded from this review paper since the effect of these mutations is very difficult to interpret based on the structures of PheOH. Most of these mutations putatively lead to truncated versions of hPheOH and no detectable level of PheOH enzyme is expressed.

One of the two deletion mutants that have been expressed *in vitro*, the 3-bp in-frame deletion of I94, was expressed using human colon adenocarcinoma cells (SW613-12A1) (40). This mutant showed 27% residual activity and markedly reduced binding affinity for L-Phe substrate. Ile94 is located in the middle of α-helix Rα2. A deletion of this hydrophobic isoleucine would disturb the helix packing by giving a one residue shift in the hydrophobic versus hydrophilic residues of helix Rα2.

TABLE 2
Deletion and Insertion Mutations Identified in the PheOH Gene

cDNA mutation	PheOH mutation	Structural contacts/comments	Reference
delE1/E2	delE1/E2	Deletion of residues 1–57 in RD. Some active PheOH protein may be expressed.	(87)
(115–117)del3bp	delF39	In center of RD in cluster of aromatic residues. Deletion of Phe disturbs packing of RD hydrophobic core.	(39)
GAAdel	E43/E44del	Deletion of both Glu disrupts hydrogen bonds in RD.	TC
(131–133)del3bp	K42–V45delAAG	In loop between Rβ1 and Rα1. Deletion disturbs RD packing.	(36)
del3bpTTT	delF55	At center of RD. Deletion disturbs hydrophobic packing at RD core.	(77)
Gg/Aa	delE2/IVS2nt1	Deletion of residues 21 to 56 in RD. Some active PheOH protein may be expressed.	(88)
delE3	delE3	Deletion of residues 57 to 118 of RD.	(89)
(208–210)del3bp	delS70	At end of Rβ2. H bonds keep loop in place for Rβ2 to pack properly.	TC
(280–282)del3bp	delI94	In middle of Rα2. Deletion disturbs helix formation.	(40)
(442–706)del	442–706delE5/E6	Deletion of residues 147 to 235 in CD. May result in very unstable or little PheOH protein expressed.	TC
(895–897)del3bp	delF299	In a hydrophobic area of CD towards end of Cα8. Deletion will disturb CD fold.	(52)
(967–969)delACA	delT323	In Cα9. H bonds to Glu319 carbonyl oxygen. Deletion disturbs CD fold.	(52)
(1066 → 1071)ins6bp	Y356/C357ins6bp	At surface of CD. Room for two extra residues with only small changes to secondary structure in this loop region.	(36)
(1090 → 1092)del3bp	delL364	No secondary structure in region. Located between two prolines (362 and 366). Residue may be necessary to position prolines for correct fold of CD.	(33)
(1090 → 1104)	delLLPLEdel15bp	Deletion of large portion of CD. Wrong CD fold affects TD and CD.	TC
del15bp 1315nt1g → a	IVS12 + 1g → a	Mutation in intron 12 of PheOH. Leads to expression of truncated PheOH missing last 52 residues.	(42)

The second PKU deletion mutation is delL364. Leu364 is positioned between two proline residues (Pro362 and Pro366) and may be a necessary spacer residue for the prolines. Deletion of this residue would result in improper positioning of the rigid prolines and thus perturb the entire catalytic domain fold, corresponding to no detectable levels of enzymatic activity as found upon expression of the mutant in COS cells (33).

The splicing mutation IVS12 + 1g → a in intron 12 of the *PAH* gene is the most prevalent PKU allele among Caucasians (41). The clinical and metabolic phenotype of homozygotes with the mutation is a severe ("classical") form of PKU. The mutation leads to the synthesis of a truncated form of PheOH lacking the last 52 amino acids in the C-terminal tetramerization domain. Expression studies in eukaryotic systems indicate that the deletion abolishes PheOH activity in the cell as a result of protein instability (42).

SUMMARY

The recently solved crystal structures of human and rat PheOH have proven to be valuable in the interpretation of many of the mutations that cause PKU. However, the effects of some of the reported mutations cannot be predicted based on either the PheOH structure or the *PAH* gene sequence and thus *in vitro* and *in vivo* expression analyses are important tools in determining the details of PKU. In order to develop the complete picture of the mutations that lead to PKU, a great deal of information still needs to be collected on both the frequency of different PKU mutations around the world and a continuing effort on the expression analyses of mutations in the *PAH* gene. The frequency of mutations as calculated by the occurrence of CpG dinucleotides (by the program MUTPRED) (15) and actual occurrence of mutations in the cases reported at the PKU database do seem to match in most cases (Fig. 3B),

like for Arg158, Arg252, Arg261, Arg408, and Asp415. These are all sites of CpG dinucleotides in the *PAH* gene sequence and are thus prone to mutability (15). In other cases, residues that originate from non-CpG dinucleotides, like Gly46, Leu48, Ile65, and Val388, still have a high frequency of PKU mutation and some are also associated with a severe PKU phenotype.

ACKNOWLEDGMENTS

The authors thank Dr. Charles R. Scriver and co-workers at DeBelle Laboratory for Biochemical Genetics, Montreal Children's Hospital, Montreal, Canada, for supplying us with the PAH Mutation Analysis Consortium Newsletter for the Spring of 1998, which was used extensively in the preparation of this paper. Supplying information was also taken from the Phenylalanine Hydroxylase Mutation Analysis Consortium Database at http://www.mcgill.ca/pahdb. Atomic coordinates of the phenylalanine hydroxylase structures are available at the Protein Data Bank (PDB) at http://www.rcsb.org/pdb, using the id codes 1pah, 2pah, 1phz, and 2phm. Professor Torgeir Flatmark, Department of Biochemistry and Molecular Biology, University of Bergen, Norway, and Albert E. Beuscher are thanked for valuable help on discussions and preparation of this manuscript.

REFERENCES

1. Kaufman S. The phenylalanine hydroxylating system. *Adv Enzymol* **70:**77–264, 1993.
2. Kappock TJ, Caradonna JP. Pterin-dependent amino acid hydroxylases. *Chem Rev* **96:**2659–2756, 1996.
3. Hufton SE, Jennings IG, Cotton RGH. Structure and function of the aromatic amino acid hydroxylases. *Biochem J* **311:**353–366, 1995.
4. Waters PJ, Parniak MA, Nowacki P, Scriver CR. In vitro expression analysis of mutations in phenylalanine hydroxylase: Linking genotype to phenotype and structure to function. *Hum Mutat* **11:**4–17, 1998.
5. Scriver CR, Kaufman S, Eisensmith RC, Woo SLC. The hyperphenylalaninemias. In The Metabolic and Molecular Bases of Inherited Disease, 7th ed. (Scriver CR, Beaudet AL, Sly WS, Valle D, Eds.). New York: McGraw-Hill, pp 1015–1075, 1995.
6. Hoang L, Byck S, Prevost L, Scriver CR. PAH Mutation Analysis Consortium Database: A database for disease-producing and other allelic variations at the human PAH locus. *Nucleic Acids Res* **24:**127–131, 1996.
7. Nowacki P, Byck S, Prevost L, Scriver CS. The PAH mutation analysis consortium database: Update 1996. *Nucleic Acids Res* **25:**139–142, 1997.
8. Nowacki P, Byck S, Prevost L, Scriver CR. PAH Mutation Analysis Consortium Database: 1997. Prototype for relational locus-specific mutation databases. *Nucleic Acids Res* **26:**220–225, 1998.
9. Erlandsen H, Fusetti F, Martinez A, Hough E, Flatmark T, Stevens RC. Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria. *Nature Struct Biol* **4:**995–1000, 1997.
10. Fusetti F, Erlandsen H, Flatmark T, Stevens RC. Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. *J Biol Chem* **273:**16962–16967, 1998.
11. Kobe B, Jennings IG, House CM, Michell BJ, Goodwill KE, Santarsiero BD, Stevens RC, Cotton RGH, Kemp BE. Structural basis of autoregulation of phenylalanine hydroxylase. *Nature Struct Biol* **6:**442–448, 1999.
12. Goodwill KE, Sabatier C, Stevens RC. Crystal structure of tyrosine hydroxylase with bound cofactor analogue and iron at 2.3A resolution: Self-hydroxylation of Phe300 and the pterin-binding site. *Biochemistry* **37:**13437–13445, 1998.
13. Martinez A, Knappskog PM, Olafdottir S, Døskeland AP, Eiken HG, Svebak RM, Bozzini M, Apold J, Flatmark T. Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. *Biochem J* **306:**589–597, 1995.
14. Flatmark T, Knappskog PM, E, Martinez A. Molecular characterization of disease related mutant forms of human phenylalanine hydroxylase and tyrosine hydroxylase. In Chemistry and Biology of Pteridines and Folates (Pfleiderer W, Rokos H, Eds.). London: Blackwell Science, pp 503–508, 1997.
15. Cooper DN, Krawczak M. The mutational spectrum of single base-pair substitutions causing human genetic disease; patterns and predictions. *Hum Genet* **85:**55–74, 1990.
16. Jennings IG, Kemp BE, Cotton RGH. Localization of cofactor binding sites with monoclonal anti-idiotype antibodies: phenylalanine hydroxylase. *Proc Natl Acad Sci USA* **88:**5734–5738, 1991.
17. Knappskog PM, Eiken HG, Martinez A, Bruland O, Apold J, Flatmark T. PKU mutation (D143G) associated with an apparent high residual enzyme activity: Expression of a kinetic variant form of phenylalanine hydroxylase in three different systems. *Hum Mutat* **8:**236–246, 1996.
18. Bjørgo E, Knappskog PM, Martinez A, Stevens RC, Flatmark T. Partial characterization and three-dimensional-structural localization of eight mutations in exon 7 of the human phenylalanine hydroxylase gene associated with phenylketonuria. *Eur J Biochem* **257:**1–10, 1998.
19. Okano Y, Eisensmith RC, Guttler F, Lichter-Konecki U, Konecki DS, Trefz FK, Dasovich M, Wang T, Henriksen K, Lou H, *et al.* Molecular basis of phenotypic heterogeneity in phenylketonuria. *N Engl J Med* **324:**1232–1238, 1991.
20. Knappskog PM, Eiken HG, Martinez A, Olafsdottir S, Haavik J, Flatmark T, Apold J. Expression of wild-type and mutant forms of human phenylalanine hydroxylase in E. coli. In Chemistry and Biology of Pteridines and Folates (Ayling JE, *et al.*, Eds.). New York: Plenum Press, pp 59–62, 1993.
21. Baric I, Mardesic D, Sarnavoka V, Lichter-Konecki U, Konecki DS, Trefz FK. Geographical distribution of the P281L mutation at the phenylalanine hydroxylase locus: Possible origin in southeastern Europe. *J Inherit Metab Dis* **7:**376–377, 1994.
22. Dworniczak B, Grudda K, Stumper J, Bartholome K, Aulehla-Scholz C, Horst J. Phenylalanine hydroxylase gene:

Novel missense mutation in exon 7 causing severe phenylketonuria. *Genomics* **9:**193–199, 1991.

23. Okano Y, Wang T, Eisensmith RC, Longhi R, Riva E, Giovannini M, Cerone R, Romano C, Woo SL. Phenylketonuria missense mutations in the Mediterranean. *Genomics* **9:**96–103, 1991.
24. Benit P, Rey F, Melle D, Munnich A, Rey J. Five novel missense mutations of the phenylalanine hydroxylase gene in phenylketonuria. *Hum Mutat* **4:**229–231, 1994.
25. Guldberg P, Levy HL, Hanley WB, Koch R, Matalon R, Rouse BM, Trefz F, de la Cruz F, Henriksen KF, Guttler F. Phenylalanine hydroxylase gene mutations in the United States: Report from the Maternal PKU Collaborative Study. *Am J Hum Genet* **59:**84–94, 1996.
26. Knappskog PM, Eiken HG, Martinez A, Flatmark T, Apold J. The PKU mutation S349P causes complete loss of catalytic activity in the recombinant phenylalanine hydroxylase enzyme. *Hum Genet* **95:**171–173, 1995.
27. De Lucca M, Perez B, Desviat LR, Ugarte M. Molecular basis of phenylketonuria in Venezuela: Presence of two novel null mutations. *Hum Mutat* **11:**354–359, 1998.
28. Eiken HG, Knappskog PM, Boman H, Thune KS, Kaada G, Motzfeldt K, Apold J. Relative frequency, heterogeneity, and geographic clustering of PKU mutations in Norway. *Eur J Hum Genet* **4:**205–213, 1996.
29. John SWM, Scriver CR, Laframboise R, Rozen R. In vitro and in vivo correlations for I65T and M1V mutations at the phenylalanine hydroxylase locus. *Hum Mutat* **1:**147–153, 1992.
30. Waters PJ, Parniak MA, Hewson AS, Scriver CR. Alterations in protein aggregation and degradation due to mild and severe missense mutations (A104D, R157N) in the human phenylalanine hydroxylase gene (PAH). *Hum Mutat* **12:**344–354, 1998.
31. Okano Y, Wang T, Eisensmith RC, Steinmann B, Gitzelmann R, Woo SL. Missense mutations associated with RFLP haplotypes 1 and 4 of the human phenylalanine hydroxylase gene. *Am J Hum Genet* **46:**18–25, 1990.
32. Wang T, Okano Y, Eisensmith RC, Lo WH, Huang SZ, Zeng YT, Yuan LF, Liu SR, Woo SL. Missense mutations prevalent in Orientals with phenylketonuria: Molecular characterization and clinical implications. *Genomics* **10:**449–456, 1991.
33. Svensson E, von Dobeln U, Eisensmith RC, Hagenfeldt L, Woo SL. Relation between genotype and phenotype in Swedish phenylketonuria and hyperphenylalaninemia patients. *Eur J Pediatr* **152:**132–139, 1993.
34. DiLella AG, Marvit J, Brayton K, Woo SL. An amino-acid substitution involved in phenylketonuria is in linkage disequilibrium with DNA haplotype 2. *Nature* **327:**333–336, 1987.
35. Svensson E, Eisensmith RC, Dworniczak B, von Dobeln U, Hagenfeldt L, Horst J, Woo SL. Two missense mutations causing mild hyperphenylalaninemia associated with DNA haplotype 12. *Hum Mutat* **1:**129–137, 1992.
36. Tyfield LA, Stephenson A, Cockburn F, Harvie A, Bidwell JL, Wood NA, Pilz DT, Harper P, Smith I. Sequence variation at the phenylalanine hydroxylase gene in the British Isles. *Am J Hum Genet* **60:**388–396, 1997.
37. Perez B, Desviat LR, De Lucca M, Ugarte M. Spectrum and origin of phenylketonuria mutations in Spain. *Acta Paediatr Suppl* **407:**34–36, 1994.
38. Mallolas J, Campistol J, Lambruschini N, Vilaseca MA, Cambra FJ, Estivill X, Mila M. Two novel mutations in exon 11 of the PAH gene (V1163del TG and P362T) associated with classic phenylketonuria and mild phenylketonuria. *Hum Mutat* **11:**482, 1998.
39. Guldberg P, Henriksen KF, Guttler F. Molecular analysis of phenylketonuria in Denmark: 99% of the mutations detected by denaturing gradient gel electrophoresis. *Genomics* **17:**141–146, 1993.
40. Caillaud C, Lyonnet S, Rey F, Melle D, Frebourg T, Berthelon M, Vilarinho L, Vaz Osorio R, Rey J, Munnich A. A 3-base pair in-frame deletion of the phenylalanine hydroxylase gene results in a kinetic variant of phenylketonuria. *J Biol Chem* **266:**9351–9354, 1991.
41. DiLella AG, Marvit J, Lidsky AS, Guttler F, Woo SLC. Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. *Nature* **322:**799–803, 1986.
42. Marvit J, DiLella AG, Brayton K, Ledley FD, Robson KJ, Woo SL. GT to AT transition at a splice donor site causes skipping of the preceding exon in phenylketonuria. *Nucleic Acids Res* **15:**5613–5628, 1987.
43. Kabsch W, Sander C. Dictionary of protein secondary structure:pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **22:**2577–2637, 1983.
44. Kayaalp E, Treacy E, Waters PJ, Byck S, Nowacki P, Scriver CR. Human phenylalanine hydroxylase mutations and hyperphenylalaninemia phenotypes: A metanalysis of genotype-phenotype correlations. *Am J Hum Genet* **61:**1309–1317, 1997.
45. Forrest SM, Dahl HH, Howells DW, Dianzani I, Cotton RG. Mutation detection in phenylketonuria by using chemical cleavage of mismatch: Importance of using probes from both normal and patient samples. *Am J Hum Genet* **49:**175–183, 1991.
46. Carter KC, Rozen R, Byck S, Scriver CR. Novel mutations and heterogeneity of the phenylalanine hydroxylase (*PAH*) gene on Quebec PKU chromosomes. *Am J Hum Genet* **57**(Suppl):A161 (Abstract), 915, 1995.
47. Guldberg P, Henriksen KF, Thony B, Blau N, Guttler F. Molecular heterogeneity of nonphenylketonuria hyperphenylalaninemia in 25 Danish patients. *Genomics* **21:**453–455, 1994.
48. Lichter-Konecki U, Schlotter M, Trefz FK, Konecki DS. Identification of the new mutations at the phenylalanine hydroxylase gene locus. Vth International Congress Inborn Errors of Metabolism W4.4, 1990 (abstract).
49. Li J, Eisensmith RC, Wang T, Lo WH, Huang SZ, Zeng YT, Yuan LF, Liu SR, Woo SL. Phenylketonuria in China: Identification and characterization of three novel nucleotide substitutions in the human phenylalanine hydroxylase gene. *Hum Mutat* **3:**312–314, 1994.
50. Bosco P, Cali F, Meli C, Mollica F, Zammarchi E, Cerone R, Vanni C, Palillo L, Greco D, Romano V. Eight new mutations of the phenylalanine hydroxylase gene in Italian patients with hyperphenylalaninemia. *Hum Mutat* **11:**240–243, 1998.
51. Horst J, Eigel A, Auleha-Scholz C, Kalaydjieva L, Zygulska

M, Kunert E, Dworniczak B. Molecular basis of phenylketonuria: Report of an extensive study of various Caucasian populations. *Am J Hum Genet* **49**(Suppl): A2302, 1991.

52. Eisensmith RC, Martinez D, Kuzmin A, Goltsov A, Woo SLC. Molecular basis of phenylketonuria in a heterogeneous US population (Abstract). *Am J Hum Genet* **57:**A163 (926), 1995.
53. Perez B, Desviat LR, Ugarte M. Analysis of the phenylalanine hydroxylase gene in the Spanish population: Mutation profile and association with intragenic polymorphic markers. *Am J Hum Genet* **60:**95–102, 1997.
54. Ramus SJ, Treacy EP, Cotton RG. Characterization of phenylalanine hydroxylase alleles in untreated phenylketonuria patients from Victoria, Australia: Origin of alleles and haplotypes. *Am J Hum Genet* **56:**1034–1041, 1995.
55. Li J, Eisensmith RC, Wang T, Lo WH, Huang SZ, Zeng YT, Yuan LF, Liu SR, Woo SL. Identification of three novel missense PKU mutations among Chinese. *Genomics* **13:**894–895, 1992.
56. Takarada Y, Kalanin J, Yamashita K, Ohtsuka N, Kagawa S, Matsuoka A. Phenylketonuria mutant alleles in different populations: Missense mutation in exon 7 of phenylalanine hydroxylase gene. *Clin Chem* **39:**2354–2355, 1993.
57. Shirahase W, Shimada AM. Genetic study on Japanese classical phenylketonuria (family analysis by PCR-SSCP analysis). *Acta Pediatr Japan* **96:**939–945, 1992.
58. Goltsov A, Kouzmine A, Eisensmith RC, Effat L, Tempamy S, Rushdi S, Abdel-Meguid N, Woo SLC. Molecular basis of phenylketonuria in Egypt (abstract). *Am J Hum Genet* **55:** 1288, 1994.
59. Kuzmin AI, Eisensmith RC, Goltsov AA, Sergeeva NA, Schwartz EI, Woo SL. Complete spectrum of PAH mutations in Tataria: Presence of Slavic, Turkic and Scandinavian mutations. *Eur J Hum Genet* **3:**246–255, 1995.
60. Zschocke J, Graham CA, Stewart FJ, Carson DJ, Nevin NC. Non-phenylketonuria hyperphenylalaninaemia in Northern Ireland: Frequent mutation allows screening and early diagnosis. *Hum Mutat* **4:**114–118, 1994.
61. van der Sijs-Bos CJ, Diepstraten CM, Juyn JA, Plaisier M, Giltay JC, van Spronsen FJ, Smit GP, Berger R, Smeitink JA, Poll-The BT, Ploos van Amstel JK. Phenylketonuria in The Netherlands: 93% of the mutations are detected by single-strand conformation analysis. *Hum Hered* **46:**185–190, 1996.
62. Hashem N, Bosco P, Chiavetta V, Cali F, Ceratto N, Romano V. Preliminary studies on the molecular basis of hyperphenylalaninemia in Egypt. *Hum Genet* **98:**3–6, 1996.
63. Popescu A, Andrian T, Guttler F, Guldberg P. Genotype-phenotype correlation in 11 Romanian PKU families. *J Inherit Metab Dis* **17:**374–375, 1994.
64. Dianzani I, Giannattasio S, de Sanctis L, Alliaudi C, Lattanzio P, Vici CD, Burlina A, Burroni M, Sebastio G, Carnevale F, *et al.* Characterization of phenylketonuria alleles in the Italian population. *Eur J Hum Genet* **3:**294–302, 1995.
65. Dworniczak B, Kalaydjieva L, Pankoke S, Aulehla-Scholz C, Allen G, Horst J. Analysis of exon 7 of the human phenylalanine hydroxylase gene: A mutation hot spot? *Hum Mutat* **1:**138–146, 1992.
66. Zschocke J, Graham CA, Carson DJ, Nevin NC. Phenylketonuria mutation analysis in Northern Ireland: A rapid stepwise approach. *Am J Hum Genet* **57:**1311–1317, 1995.
67. Okano Y, Hase Y, Lee DH, Takada G, Shigematsu Y, Oura T, Isshiki G. Molecular and population genetics of phenylketonuria in Orientals: Correlation between phenotype and genotype. *J Inherit Metab Dis* **17:**156–159, 1994.
68. Perez B, Desviat LR, Ugarte M. Expression analysis of mutation P244L, which causes mild hyperphenylalaninemia. *Hum Mutat* **5:**188–190, 1995.
69. Caillaud C, Vilarinho L, Vilarinho A, Rey F, Berthelon M, Santos R, Lyonnet S, Briard ML, Osorio RV, Rey J, *et al.* Linkage disequilibrium between phenylketonuria and RFLP haplotype 1 at the phenylalanine hydroxylase locus in Portugal. *Hum Genet* **89:**69–72, 1992.
70. Shirahase W, Oya N, Shimada M. A new single base substitution in a Japanese phenylketonuria (PKU) patient. *Brain Dev* **13:**283–284, 1991.
71. Leandro P, Rivera I, Ribeiro V, Tavares de Almeida I, Lechner MC. Sequencing analysis of PAH genomic DNA reveals 4 novel mutations affecting exons 7 and 11 in a Portugese PKU population. Abstr. SSIEM, Manchester, 1993.
72. Kleiman S, Li J, Schwartz G, Eisensmith RC, Woo SL, Shiloh Y. Inactivation of phenylalanine hydroxylase by a missense mutation, R270S, in a Palestinian kinship with phenylketonuria. *Hum Mol Genet* **2:**605–606, 1993.
73. Melle D, Verelst P, Rey F, Berthelon M, Francois B, Munnich A, Lyonnet S. Two distinct mutations at a single BamHI site in phenylketonuria. *J Med Genet* **28:**38–40, 1991.
74. Goebel-Schreiner B, Schreiner R. Identification of a new missense mutation in Japanese phenylketonuric patients. *J Inherit Metab Dis* **16:**950–956, 1993.
75. Labrune P, Melle D, Rey F, Berthelon M, Caillaud C, Rey J, Munnich A, Lyonnet S. Single-strand conformation polymorphism for detection of mutations and base substitutions in phenylketonuria. *Am J Hum Genet* **48:**1115–1120, 1991.
76. Takahashi K, Masamune A, Kure S, Matsubara Y, Narisawa K. Ectopic transcription: An application to the analysis of splicing errors in phenylalanine hydroxylase mRNA. *Acta Paediatr Suppl* **407:**45–46, 1994.
77. Zygulska M, Eigel A, Pietrzyk JJ, Horst J. Phenylketonuria in southern Poland: A new splice mutation in intron 9 at the PAH locus. *Hum Mutat* **4:**297–299, 1994.
78. Eisensmith RC, Woo SL. Molecular basis of phenylketonuria and related hyperphenylalaninemeias: Mutations and polymorphisms in the human phenylalanine gene. *Hum Mutat* **1:**13–23, 1992.
79. Kalaydjieva L, Dworniczak B, Kremensky I, Koprivarova K, Radeva B, Milusheva R, Aulehla-Scholz C, Horst J. Heterogeneity of mutations in Bulgarian phenylketonuria haplotype 1 and 4 alleles. *Clin Genet* **41:**123–128, 1992.
80. Economou-Petersen E, Henriksen KF, Guldberg P, Guttler F. Molecular basis for nonphenylketonuria hyperphenylalaninemia. *Genomics* **14:**1–5, 1992.
81. Rozen R, Mascisch A, Lambert M, Laframboise R, Scriver CR. Mutation profiles of phenylketonuria in Quebec populations: Evidence of stratification and novel mutations. *Am J Hum Genet* **55:**321–326, 1994.
82. Lichter-Konecki U, Konecki DS, DiLella AG, Brayton K, Marvit J, Hahn TM, Trefz FK, Woo SL. Phenylalanine hy-

droxylase deficiency caused by a single base substitution in an exon of the human phenylalanine hydroxylase gene. *Biochemistry* **27:**2881–2885, 1988.

83. Weinstein M, Eisensmith RC, Abadie V, Avigad S, Lyonnet S, Schwartz G, Munnich A, Woo SL, Shiloh Y. A missense mutation, S349P, completely inactivates phenylalanine hydroxylase in north African Jews with phenylketonuria. *Hum Genet* **90:**645–649, 1993.
84. Desviat LR, Perez B, De Lucca M, Cornejo V, Schmidt B, Ugarte M. Evidence in Latin America of recurrence of V388M, a phenylketonuria mutation with high in vitro residual activity. *Am J Hum Genet* **57:**337–342, 1995.
85. Desviat LR, Perez B, Ugarte M. Molecular basis of non-PKU hyperphenylalaninaemia in Spain: Prevalence of A403V, a mutation with high residual activity. *J Inherit Metab Dis* **19:**227–230, 1996.
86. Wang T, Okano Y, Eisensmith RC, Harvey ML, Wilson HYL, Huang S-Z, Zeng Y-T, Yuan L-F, Furuyama J-I, Oura T, Sommer SS, Woo SLC. Founder effect of a prevalent phenylketonuria mutation in the Oriental population. *Proc Natl Acad Sci USA* **88:**2146–2150, 1991.
87. Sullivan SE, Lidsky AS, Brayton K, DiLella AG, King M, Connor M, Cockburn F, Woo SLC. Phenylalanine hydroxylase deletion mutant from a patient with classical PKU. *Am J Hum Genet* **37**(Suppl):A177, 1985.
88. Kleiman S, Bernstein J, Schwartz G, Eisensmith RC, Woo SL, Shiloh Y. A defective splice site at the phenylalanine hydroxylase gene in phenylketonuria and benign hyperphenylalaninemia among Palestinian Arabs. *Hum Mutat* **1:**340–343, 1992.
89. Avigad S, Cohen BE, Bauer S, Schwartz G, Frydman M, Woo SL, Niny Y, Shiloh Y. A single origin of phenylketonuria in Yemenite Jews. *Nature* **344:**168–170, 1990.